NASA Contractor Report 171984

National Aeronautics and Space Administration (NASA)/ American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program--1986

Volume 1

Bayliss McInnis, Editor
University of Houston--University Park
Houston, Texas

Stanley Goldstein, Editor
University Programs Office
Lyndon B. Johnson Space Center
Houston, Texas

NASA

National Aeronautics and
Space Administration

Lyndon B. Johnson Space Center
Houston, Texas

PREFACE

The 1986 Johnson Space Center (JSC) National Aeronautics and Space Admin-
istration (NASA)/American Society for Engineering Education (ASEE) Summer
Faculty Fellowship Program was conducted by the University of Houston and
JSC. The ten week program was operated under the auspices of the ASEE.
The program at JSC, as well as the programs at other NASA Centers, was
funded by the Office of University Affairs, NASA Headquarters, Washington,
D.C. The objectives of the programs, which began in 1965 at JSC and in
1964 nationally, are

a. to further the professional knowledge of qualified engineering and
    science faculty members;

b. to stimulate an exchange of ideas between participants and NASA;

c. to enrich and refresh the research and teaching activities of
    participants' institutions; and

d. to contribute to the research objectives of the NASA Centers.

Each faculty fellow spent ten weeks at JSC engaged in a research project
commensurate with his interests and background and worked in collaboration
with a NASA/JSC colleague. This document is a compilation of the final
reports on the research projects done by the faculty fellows during the
summer of 1986. Volume 1 contains sections 1 through 14, and volume 2
contains sections 15 through 30.

CONTENTS

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

University of Houston

Spectral Characterization of Martian Soil Analogues

Prepared by: David G. Agresti, Ph. D.

Academic Rank: Associate Professor of
                                        Physics

University and Department: University of Alabama
                                        at Birmingham

NASA/JSC

Directorate: Space and Life Sciences

Division: Solar System Exploration

Branch: Experimental Planetology

JSC Colleague: Richard V. Morris, Ph. D.

Date: August 15, 1986

Contract #:   NGT-44-005-803
                                        (University of Houston)

SPECTRAL CHARACTERIZATION OF MARTIAN SOIL ANALOGUES

David G. Agresti, Ph. D.
Associate Professor of Physics
University of Alabama at Birmingham
Birmingham, AL   35294

ABSTRACT

As reported by Dr. Richard Morris in March, reflectance spectra of iron (III) oxide precipitated as ultrafine (x-ray amorphous) particles, unlike ordinary fine-grained (>100 nm) hematite (α-Fe_2O_3), have significant similarities to reflectance spectra from the bright regions of Mars. During this summer's stay at JSC, I have collaborated with Dr. Morris to characterize these particles according to composition, magnetic properties, and particle-size distribution. Mossbauer, magnetic susceptibility, and optical data have been obtained for samples with a range of concentrations of iron oxide in silica gel of varying pore diameters (6, 15, and 30 nm). To analyze the Mossbauer spectra, I have enhanced a versatile fitting program (adapted during last summer's ASEE visit for the IBM-PC) to provide user-friendly screen input and theoretical models appropriate for the superparamagnetic spectra obtained.

NASA Colleague: Richard V. Morris, Ph. D.,   SN4,   X5874

INTRODUCTION

In March of this year [Morris and Lauer, 1986], Dr. Richard Morris reported that hematite formed by calcining silica gel impregnated with ferric nitrate solution provided a material with spectra similar to reflectance spectra obtained from the bright regions of Mars. The hematite has an ultrafine particle size because of the small pore size (6-30 nm) of the silica gel matrix in which it is prepared. Further work in Dr. Morris' laboratory at JSC has been aimed at producing a variety of samples of this material in gels of various pore sizes and under varying conditions of temperature of oxidation, pH, starting reagents, etc., in order to produce the material with the best match to the Mars spectrum. Other techniques are employed to further characterize the properties of these martian soil analogues and to determine the magnetic properties, chemical composition, size distribution, etc., of the particles of which this material is made.

One of the tools employed by Dr. Morris in this effort is Mossbauer spectroscopy. In this technique, a spectrum is obtained by allowing radiation emitted from a moving source of 14.4-keV gamma-rays from ^{57}Fe (about 2% of natural iron) to pass through the material (absorber) under study. This transmission spectrum is obtained in a multichannel analyzer as a plot of number of gamma-ray counts detected versus the

velocity of the source (1 mm/s corresponds to an energy shift of 4.8×10^{-8} eV) [for example, Figure 1]. The spectrum is analyzed by computer fitting a theoretical function to the data and interpreting the fitted parameters in terms of the environment of the iron atoms in the absorber material.

In the present study, spectra generally consist of a 6-peak magnetic spectrum superimposed on a 2-peak paramagnetic spectrum [Figure 2]. The simplest interpretation of the spectra observed is that the iron atoms are in two chemically distinct sites, one of which results in a magnetic field at the iron nucleus. However, it is well-known that iron oxide particles < 30nm in diameter exhibit superparamagnetism [Kundig et al., 1966]. In this report, the phenomenon of superparamagnetism will be discussed and applied to the silica-gel oxides to obtain useful information about the particle-size distribution in these samples.

In continuing with the adaptation for the IBM-PC of the Mossbauer data least-squares fitting program [Agresti et al., 1969] begun last summer [Agresti, 1985], I have attempted to provide a more user-friendly screen input and fitting models particularly suited to the Mossbauer spectra obtained on the martian soil analogues. In this report, several of these enhancements will be described and illustrated by application to spectra obtained in Dr. Morris' laboratory.

SUPERPARAMAGNETISM

As stated in the introduction, iron oxide formed in silica gel may be superparamagnetic. The samples obtained generally have Mossbauer spectra with a 6-peak magnetic component, which implies the presence of a magnetic field at the nucleus (31.15 kOe per mm/s of splitting between the outer pair of lines), and a 2-peak paramagnetic component, which implies the absence of a magnetic field, or more properly, a zero time-average magnetic field. The 2-peak splitting results from the distortion of the local environment of the iron atom from cubic symmetry. Superparamagnetic particles are so small that the magnetic anisotropy energy, which is proportional to volume, is not sufficient to maintain the domain magnetization pointing permanently in one of several possible easy directions in the crystal, and the magnetizetion flips among easy directions with a frequency related to the thermal energy, kT. The reciprocal of this frequency, the relaxation time, t_o, is proportional to the Maxwell-Boltzmann weighting factor:

$$t_o \propto \exp{(2KV/kT)}, \tag{1}$$

where K is the magnetic anisotropy constant and V is the volume of the superparamagnetic particle.

In order for a magnetic field to be observed at the nucleus (resulting in a 6-line pattern), t_o must be long compared to the time of observation, t_{obs}; a 2-line pattern

will appear when t_o is much shorter than t_{obs}. The time, t_{obs}, is necessary to establish the value of the field at the nucleus, and, from the Heisenberg uncertainty product, is equal to the nuclear level splitting resulting from the magnetic field divided by Planck's constant, $\hbar$. (For the 500 kOe fields of α -Fe$_2$O$_3$, t_{obs} is approx. equal to 2.5×10^{-8} sec). Thus, from the spectrum shown in Figure 2, our sample consists of a distribution of particle sizes, the smaller particles being associated with the doublet and the larger particles with the sextet. The area under each of these two components of the spectrum is proportional to the number of nuclei, that is, the total volume, in each size regime.

To obtain a size distribution, and also to confirm the supermagnetic nature of our samples, it is necessary to collect Mossbauer spectra over a range of temperatures. Figure 3 shows a series of Mossbauer spectra collected down to 22K on a silica gel sample supplied by Dr. Morris. These were taken by my graduate student and NASA Graduate Trainee Jeffrey Newcomb at UAB. In the figure, it is seen that there is a steady increase with temperature of the 2-peak component at the expense of the 6-peak component.

To explain this effect, we point out that Equation (1) shows that t_o depends on temperature as well as on particle size. In fact, the exponential dependence implies a fairly sharp transition as a function of temperature, for a given particle volume V, between a 6-peak and a 2-peak contribution

1-6

to the spectrum. From another point of view, for each temperature there is a transitional volume, V_t, that divides the distribution into two parts. For $V > V_t$, the particles contribute to the 6-peak component; for $V < V_t$, to the 2-peak component. Kundig et al. determined that the anisotropy constant, K, for hematite is approximately independent of temperature and gave a value of $(4.1 \pm 1) \times 10^4$ erg/cm^3. With this value and the requirement that the relaxation time, t_o, for particles of volume, V_t, be approx. $= t_{obs}$, Equation (1) may be transformed [from Eq. (8), Kundig et al., 1966] to the more convenient form,

$$V_t = [(4.7 \pm 1) \text{ nm}^3] \text{ T.} \tag{2}$$

The spectra of Figure 3 were fit to determine the relative area of the 2-peak component. Figure 4 is a graph of the results with a smooth curve drawn through the data. The curve may be understood to be proportional to the integral of the distribution, $dN(T)/dV$, which is the number of particles having volume in the range between V_t and V_t+dV, since this integral from zero K to the temperature, T, is equal to the total volume of particles with $V < V_t$. Hence, the derivative of the curve, under the assumption of constant K, gives directly a volume distribution, which may be calibrated according to Equation (2). This distribution may be converted into the desired size distribution if we assume the particles are uniform spheres, as has been done in Figure 4.

THE COMPUTER PROGRAM

The major portion of my effort this summer was devoted to enhancing the computer program [Agresti et al., 1969] used to analyze the Mossbauer data. The resulting program, along with future enhancements, will be designated "VersiFit." Last summer, the program was implemented on an IBM-PC, but, as mentioned then [Agresti, 1985], a number of modifications remained to be made. Five such enhancements will be described here in order to illustrate the range of modifications involved: These are: 1. Interactive screen input; 2. Plotting of data and fitted function; 3. Laser velocity calibration; 4. Marquardt minimization procedure; and 5. Skewed-Lorentzian peak functions.

1. <u>Interactive screen input.</u> Sample input screens are shown in Figure 5. Other input screens are provided or anticipated for input of relations among parameters, data and velocity definition, plotting requirements, etc. The basic idea is a complete break with the fixed-sequence input typical of mainframe computers. It is not only interactive, but dynamic in the sense that the user decides which information to provide through the use of the cursor controls to position the response in the correct box and through the selection of particular entry screens that contain the items required. Furthermore, the individual entry screens re-form themselves in response to earlier input, as shown by the

three screens of Figure 5. It is hoped that this more user-friendly type of input coupled with implementation on a very popular and very powerful microcomputer will ease the adoption of the program among the mineralogical community.

2. <u>Plotting</u> <u>of</u> <u>data</u> <u>and</u> <u>fitted</u> <u>function.</u> The PC used this summer is configured with a graphics printer, a monochrome monitor used for text and numerical input and output, and a high-resolution (640x200) graphics monitor used to display graphs of data and fitted function. This arrangement has proved very helpful when analyzing data. The visual display of data and function on the same graph [Figures 6,11,12] is of course much more helpful for inspecting the quality of the fit than merely noting, for example, the value of chi square (χ^2). But it is also useful for making good choices for starting values of parameters in the fit, as are plots of deviations between data and function [Figures 7,8]. With a graphics printer connected, immediate hard-copy output may be obtained for later reference or publication.

3. <u>Laser</u> <u>velocity</u> <u>calibration.</u> As described in the introduction, the spectrum is acquired in a multichannel analyzer with a moving source of radiation. The drive produces a velocity designed to be linearly proportional to channel number; thus, v_i = m x (i - 256.5), where m = velocity increment (mm/s) per channel and i ranges from 1 to 512, the number of channels. In order to obtain a precise

value for the velocity of the source, a laser is mounted parallel with the motion of the source and interference fringes are counted and stored as a function of channel. The number of fringes produced is accurately proportional to the distance covered during the period of time a channel is open, hence to the absolute value of the velocity. Figure 9 is a typical laser calibration run, associated with the data of Figures 1,2,11,12. The calibration data show that the velocity is not strictly linear, but is better represented by a "bilinear" function, $v_i = m_1 \times (i - i_0)$ for $i < i_0$, and $v_i = m_2 \times (i - i_0)$ for $i > i_0$, where m_1 and m_2 typically differ by 1%, and i_0, which corresponds to the zero-velocity channel, is generally not equal to 256.5.

4. Marquardt minimization procedure. As noted in last summer's report [Agresti, 1985], the standard non-linear least-squares fitting procedure, employing Taylor's approximation for the function, is not always successful in obtaining a minimum in χ^2, defined as

$$\chi^2 = \frac{1}{(N_d - N_p)} \sum \frac{(y_i - f_i)^2}{\sigma_i^2} \qquad (3)$$

where N_d is the number of data values (channels), N_p is the number of parameters varied in the fit, and σ_i is the standard deviation in each data value. When parameters are strongly correlated, as the peak positions of Figure 6, where the peaks strongly overlap, say in fitting with two or three overlapping doublets, then often this procedure will

produce an increasing χ^2 from iteration to iteration. A different approach [Marquardt, 1963] combines steepest descent with Taylor's approximation in such a way that the change in the parameters is neither Taylor nor steepest descent but a linear combination of the two. The linear combination is optimal in that the fit converges to a minimum in the least number of iterations. According to theory, the fit will converge, even when parameters are strongly correlated. To illustrate this, Figure 10 shows the variation in χ^2 for 6-peak fits to the data of Figure 6 with two sets of starting parameters and minimization by the Taylor or Marquardt procedure. Started close to the minimum in χ^2 of 0.966, Taylor is very sluggish compared to Marquardt; started farther away, it diverges, while Marquardt proceeds monotonically downward.

5. <u>Skewed-Lorentzian peak functions.</u> Figure 1 shows a spectrum of pure, bulk hematite. Each component peak is symmetrical in comparison to those of Figure 2. In order to accurately fit the areas, a satisfactory shape function must be supplied that will successfully reproduce the shape of the data. In the case of hematite and many other particularly well-defined crystal structures, the appropriate theoretical function is a Lorentzian, given by:

$$L(v) = \frac{\text{Area} \times 2 / (\pi W)}{1 + [2(v - v_o) / W]^2}, \qquad (4)$$

where v_o is the velocity position of the peak (more

accurately, dip) in the transmission spectrum, and W is the full-width at half maximum. Figure 11 shows a fit of the data of Figure 2 to a superposition of 8 Lorentzian peaks, with areas and widths constrained in pairs, resulting in 18 variable parameters. The value achieved for χ^2 is 8.98, and it is easy to see that the fitted function misses a great deal of the data. From the above discussion, it is evident that we must account for the asymmetry in the 6 magnetic peaks. The function I have chosen, for computational simplicity, may be termed a "skewed" Lorentzian. It has an additional parameter, δ, the "skew," and is defined (assuming $\delta > 1$) by contracting the half-width to .5 W / δ on one side of the vertical midline of the Lorentzian function and expanding the width to .5 W x δ on the other side. Figure 12 shows the fitted function obtained with the 6 magnetic peaks skewed in pairs. The final value of χ^2 has dropped to 1.35, a dramatic improvement for the addition of just 3 additional variable parameters.

In summary, improvements in the capabilities for analysis of Mossbauer spectra of martian soil analogues have been provided as a modification of an existing least-squares program, whose ease of input and variety of fitting options and models, hence VersiFit, should be of value to those in the wider scientific community who wish to employ desk-top computers in their data analysis.

ACKNOWLEDGMENTS

I wish to thank NASA and the Summer Faculty Fellowship Program for providing me with the opportunity to spend two very productive summers in a very stimulating environment. Above all, thanks to Dick Morris for much encouragement and many interesting discussions and for exposing me to an application of Mossbauer spectroscopy entirely new to me. We have started a very fine collaboration, and it will continue.

REFERENCES

1.   Agresti, D.G., M.F. Bent, and B.I. Persson, "A Versatile Computer Program for Analysis of Mossbauer Spectra," Nucl. Instr. and Methods, Vol. 72, pp. 235-236, 1969.

2.   Agresti, D.G., "Mossbauer Spectroscopy of Extraterrestrial Materials," Final Report, NASA/ASEE Summer Faculty Fellowship Program, Summer, 1985.

3.   Kundig, W., H. Bommel, G. Constabaris, and R. Lindquist, "Some Propertiesof Supported Small α-Fe_2O_3 Particles Determined with the Mossbauer Effect," Phys. Rev., Vol. 142, pp. 327-333, 1966.

4.   Marquardt, D.W., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," J. Soc. Indust. Appl. Math., Vol. 11, pp. 431-441, 1963.

5.   Morris, R.V., and H.V. Lauer, Lunar and Planetary Science Conference XVII Abstracts (Houston, March, 1986), pp. 573-574.
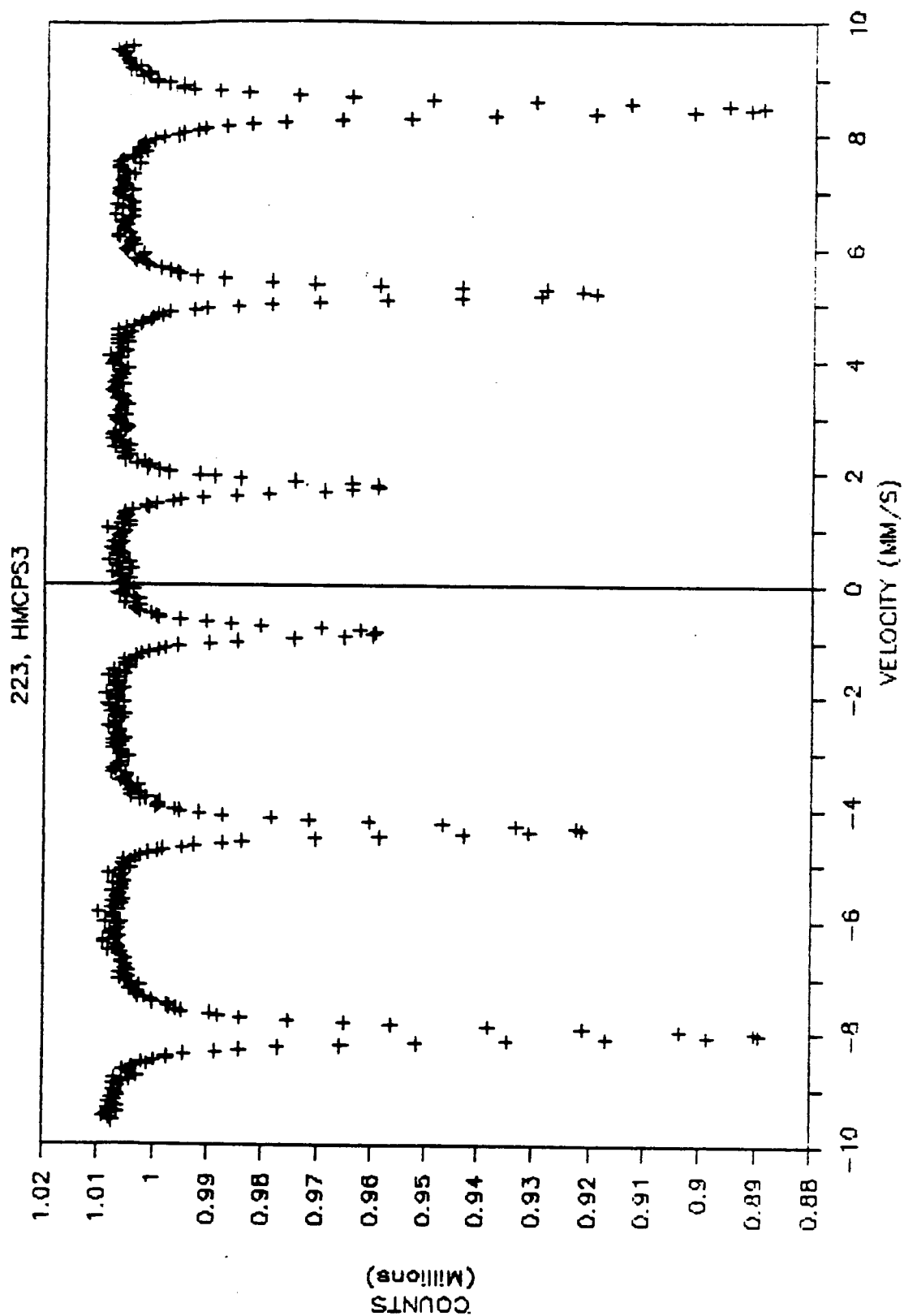
Figure 1. Mossbauer spectrum of bulk hematite.

222, SGHM2B/12/550
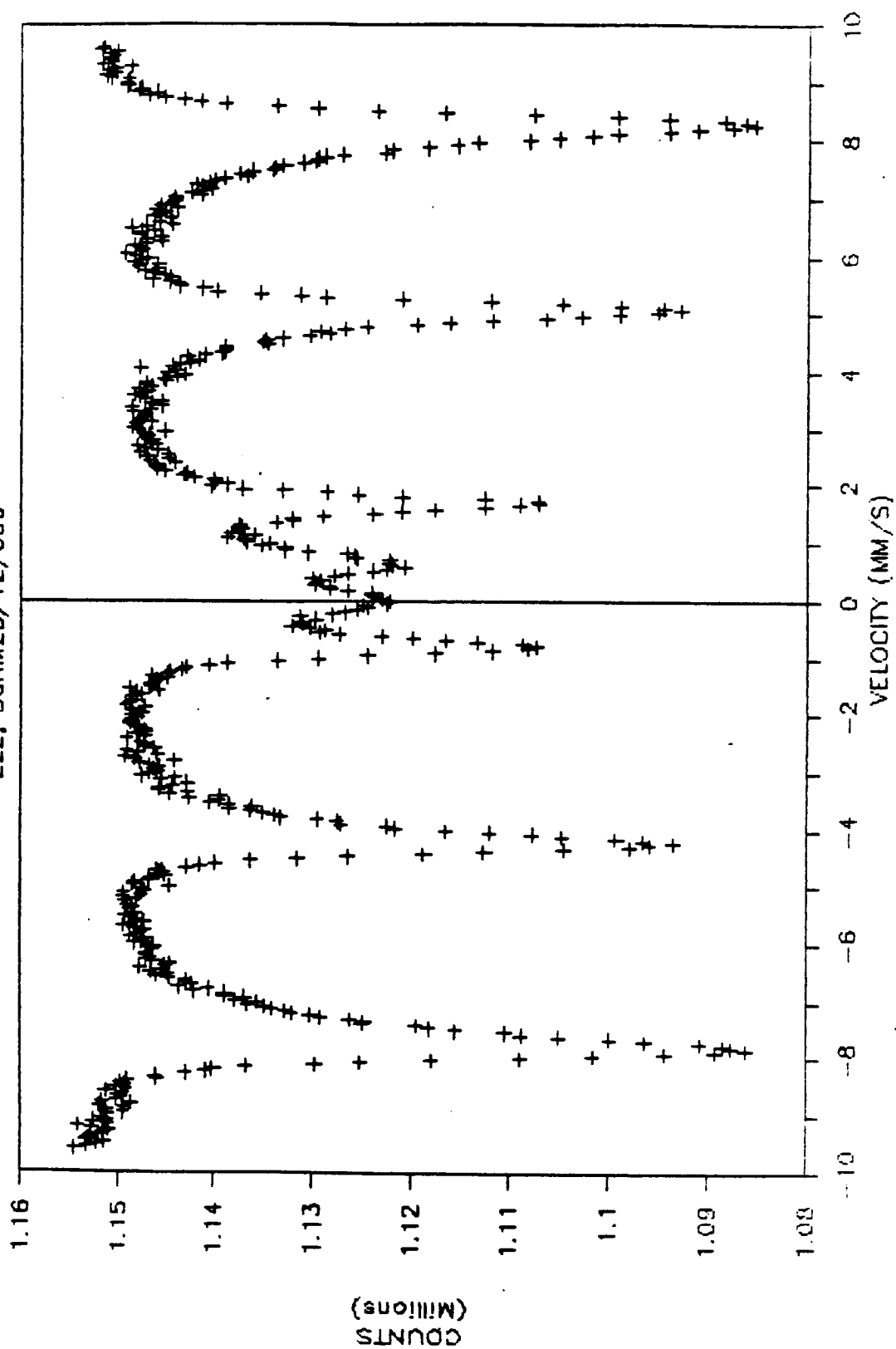
Figure 2. Room temperature Mossbauer spectrum of a sample
of Fe_2O_3 (23.7 wt%) in 6-nm silica gel.
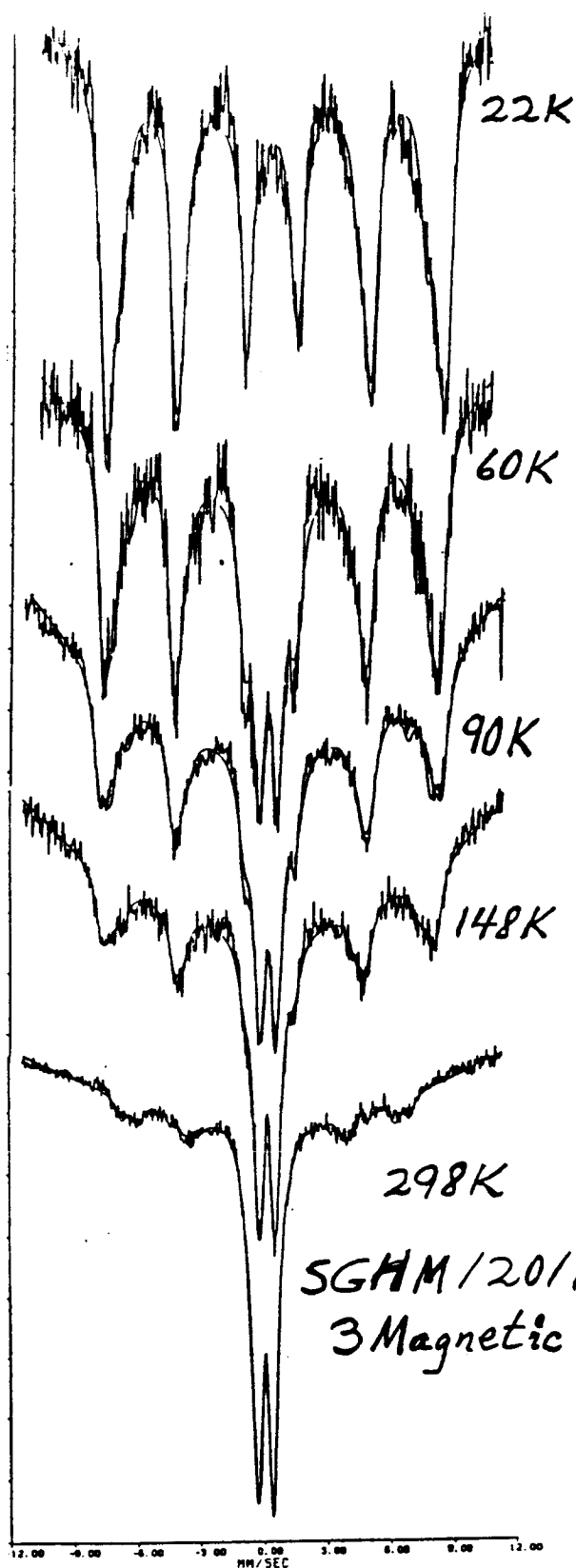
Figure 3. Mossbauer
spectra of a sample
of Fe_2O_3 (5.73 wt%)
in 6-nm silica gel,
taken at various
temperatures.
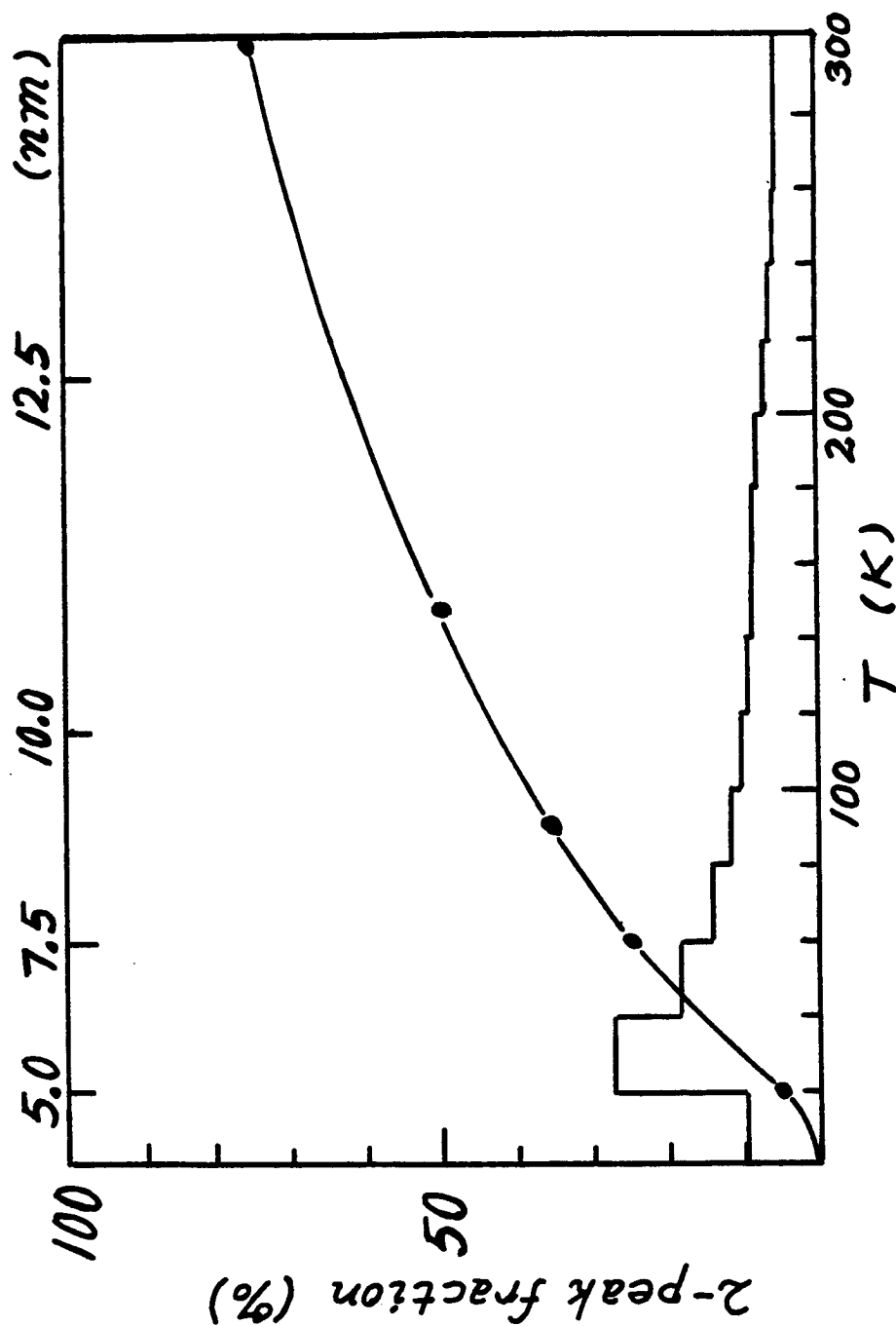
Figure 4. Fractional areas of the 2-peak paramagnetic component of the spectra of Figure 3 as function of temperature. The histogram is obtained by taking, every 20K, the vertical difference in the curve drawn through the fractional areas; it represents a volume distribution, and is calibrated with Equation (2) of the text to give a distribution of diameters of equal-volume spheres (upper scale).

```
        STARTING A NEW FIT . . .  Date:   8/20/86      Time:   1:32:18 PM

Data file name... 000206.dat          Data pts...  512   Half (1or2).... 1
Describe experiment.....
Describe the fit........
Peak to peak vel...   7.6500          Overflow...    .0   Output file... F
PSAME... F   YSAME... F   RANDOM.. F   NOSUM... T   NOFIT... T   NTCYC... 15
No. of indep peaks.....  0            No. sites in hf model.. 2

      Parameter values:
Background, B......     .5907
    Area, AREA......   7.9716    Parab background corr., GEOM...      .0000
Rel site areas..........  1.0000   1.0000
  center shifts, CS......   .2600    .4500
     quad params, QQ......  1.4800   4.8600
     mag. params, GH......   .0000    .0000
g-exc / g-gnd, GOVG.....   .0000
   equal widths, W.......   .5000
Fixed params....... DEAD    DEAD   DEAD    DEAD    DEAD    DEAD    DEAD   DEAD
```

```
        STARTING A NEW FIT . . .  Date:   8/20/86      Time:   1:26:29 PM

Data file name... 000206.dat          Data pts...  512   Half (1or2).... 1
Describe experiment.....
Describe the fit........
Peak to peak vel...   7.6500          Overflow...    .0   Output file... F
PSAME... F   YSAME... F   RANDOM.. F   NOSUM... T   NOFIT... T   NTCYC... 15
No. of indep peaks.....  2            No. sites in hf model.. 0

      Parameter values:
Background, B......     .5907
    Area, AREA......   7.9716    Parab background corr., GEOM...      .0000
Peak positions, EV......   -.1480    .6090
    rel. areas, HV......    .5120    .4880
    half-widths, WV......    .6930    .6420
 fraction Gau., FGAU....    .0000    .0000
   equal widths, W.......    .0000
Fixed params....... AREA    DEAD   DEAD    DEAD    DEAD    DEAD    DEAD   DEAD
```

```
        STARTING A NEW FIT . . .  Date:   8/20/86      Time:   1:21:24 PM

Data file name... 000206.dat          Data pts...  512   Half (1or2).... 1
Describe experiment.....
Describe the fit........
Peak to peak vel...   7.6500          Overflow...    .0   Output file... F
PSAME... F   YSAME... T   RANDOM.. F   NOSUM... T   NOFIT... T   NTCYC... 15
No. of indep peaks.....  4            No. sites in hf model.. 0

      Parameter values:
Background, B......     .5907
    Area, AREA......   7.9200    Parab background corr., GEOM...      .0000
Peak positions, EV......   -.1100    .6300   -.7700   1.6600
    rel. areas, HV......    .4000    .4000    .0500    .0500
    half-widths, WV......    .5000    .5000    .5000    .5000
 fraction Gau., FGAU....    .0000    .0000    .0000    .0000
   equal widths, W.......    .0000
Fixed params....... AREA    DEAD   DEAD    DEAD    DEAD    DEAD    DEAD   DEAD
```

Figure 5.   Examples of screen input to the program.
The screens re-form as shown in response to entries for
"No. of indep peaks" and "No. of sites in hf model."
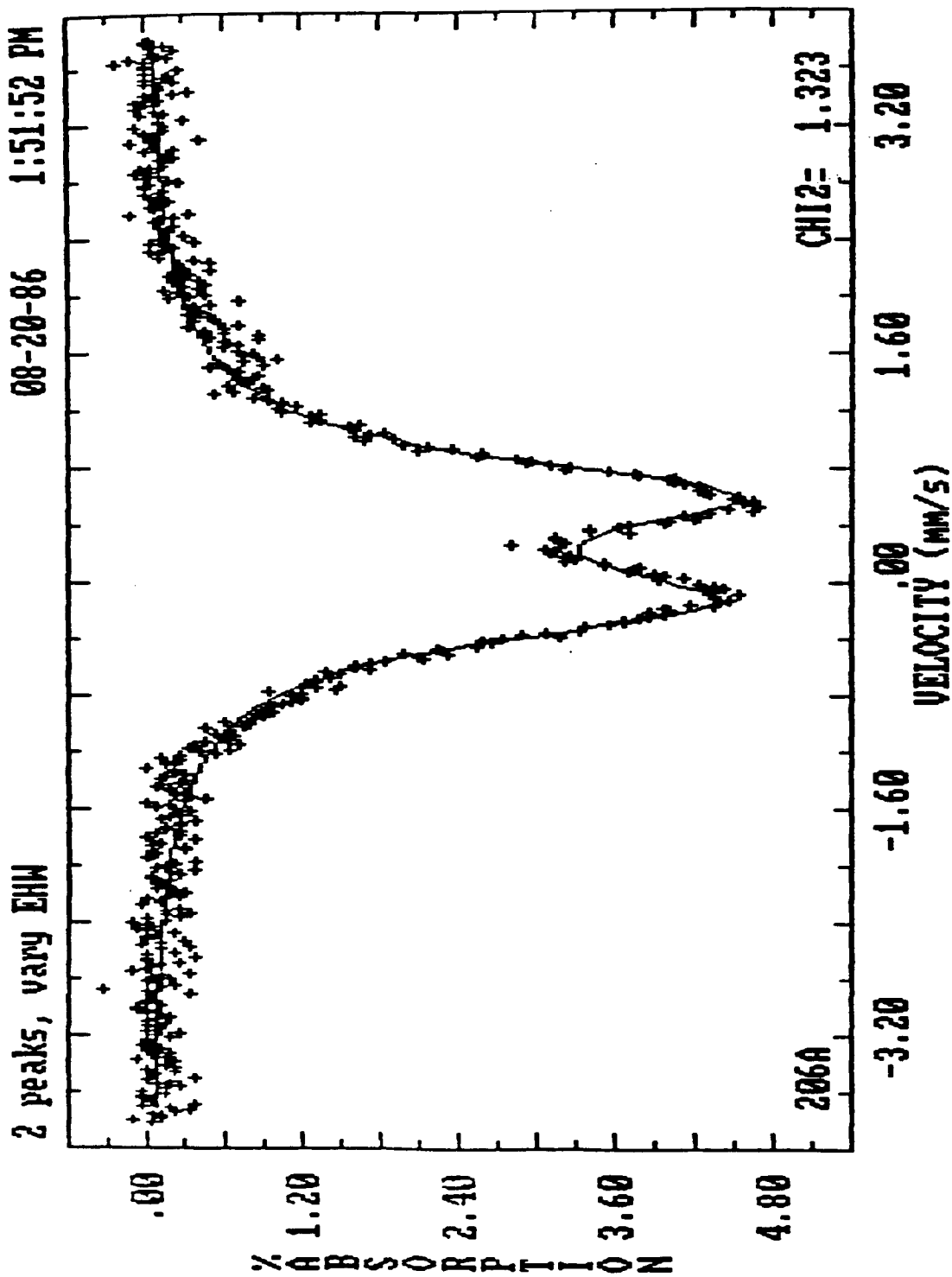
Figure 6. Plot of data and fitted function, as output by VersiFit. The data are from a sample of Fe_2O_3 (7.61 wt%) in 6-nm silica gel.

2 peaks, vary EHW   08-20-86   1:53:41 PM

REL DEV IN y: 2.00, 1.00, .00, -1.00, -2.00

VELOCITY (MM/S): -3.20, -1.60, .00, 1.60, 3.20

CHI2= 1.323
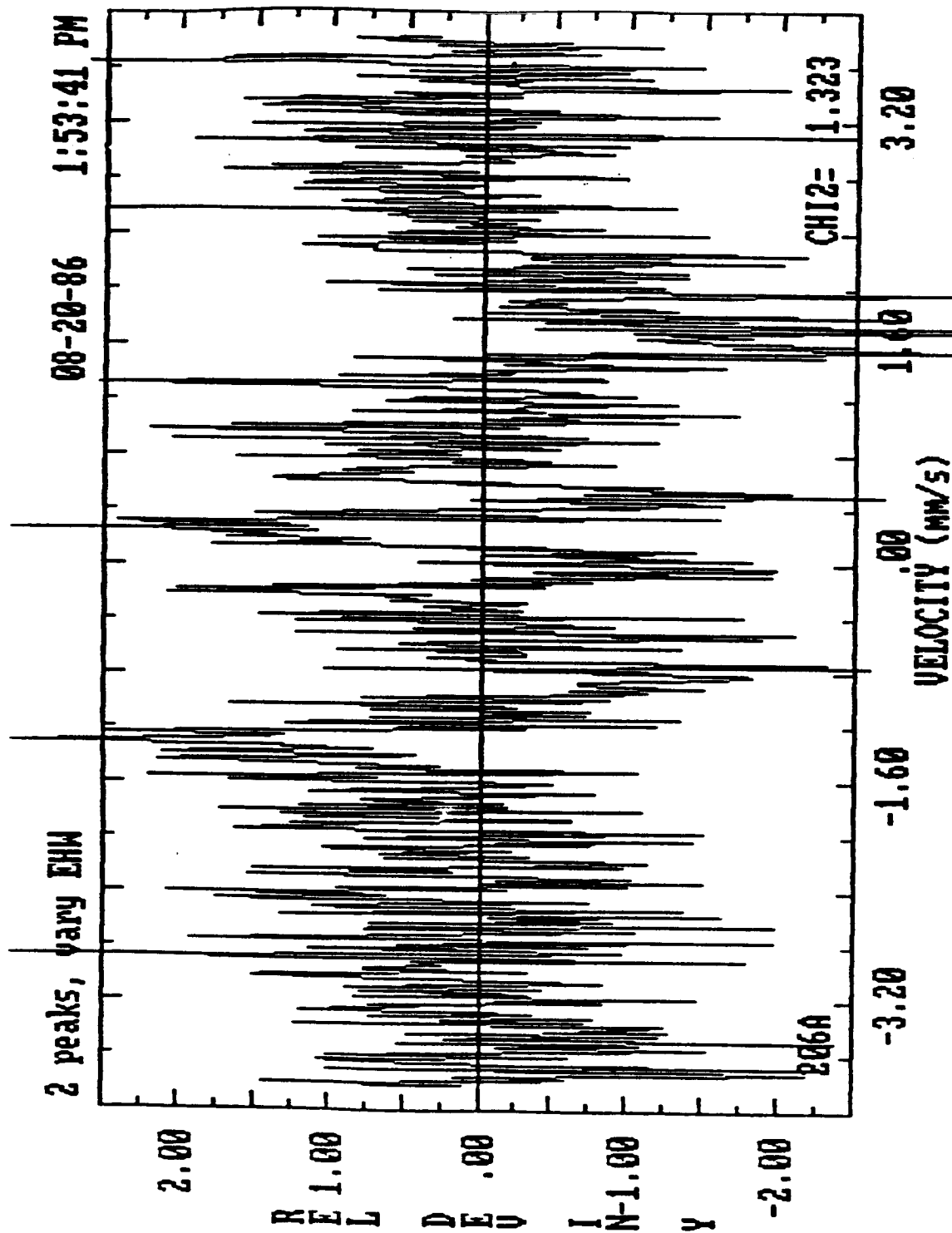
Figure 7. Relative deviation
of the data and the fitted
function of Figure 6, defined by:

$$r_i = \frac{(y_i - f_i)}{\sigma_i} = \frac{(y_i - f_i)}{y_i}$$
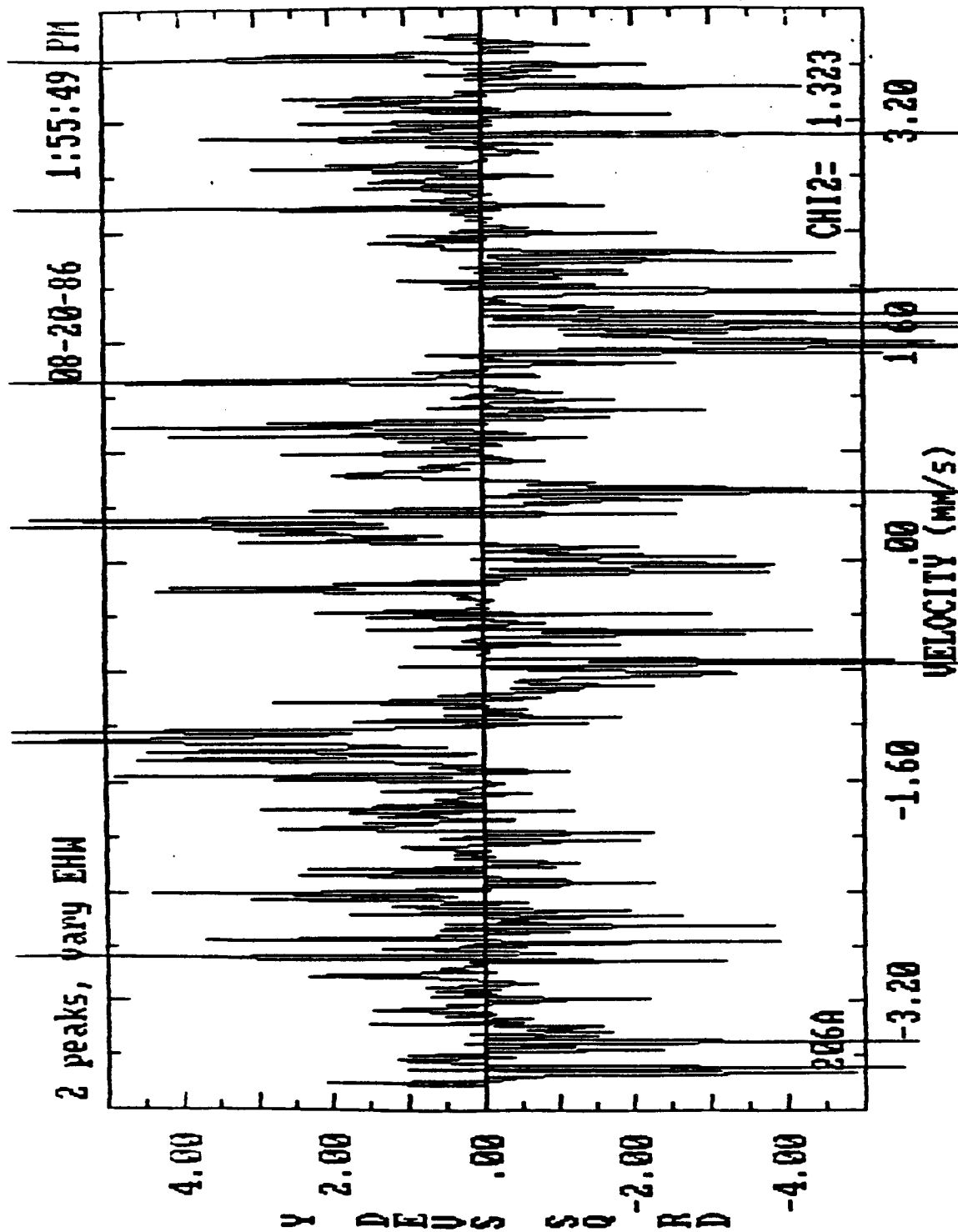
Figure 8. Squares of the values plotted in Figure 7.
These are the terms in the sum for χ^2 [Equation (3)].
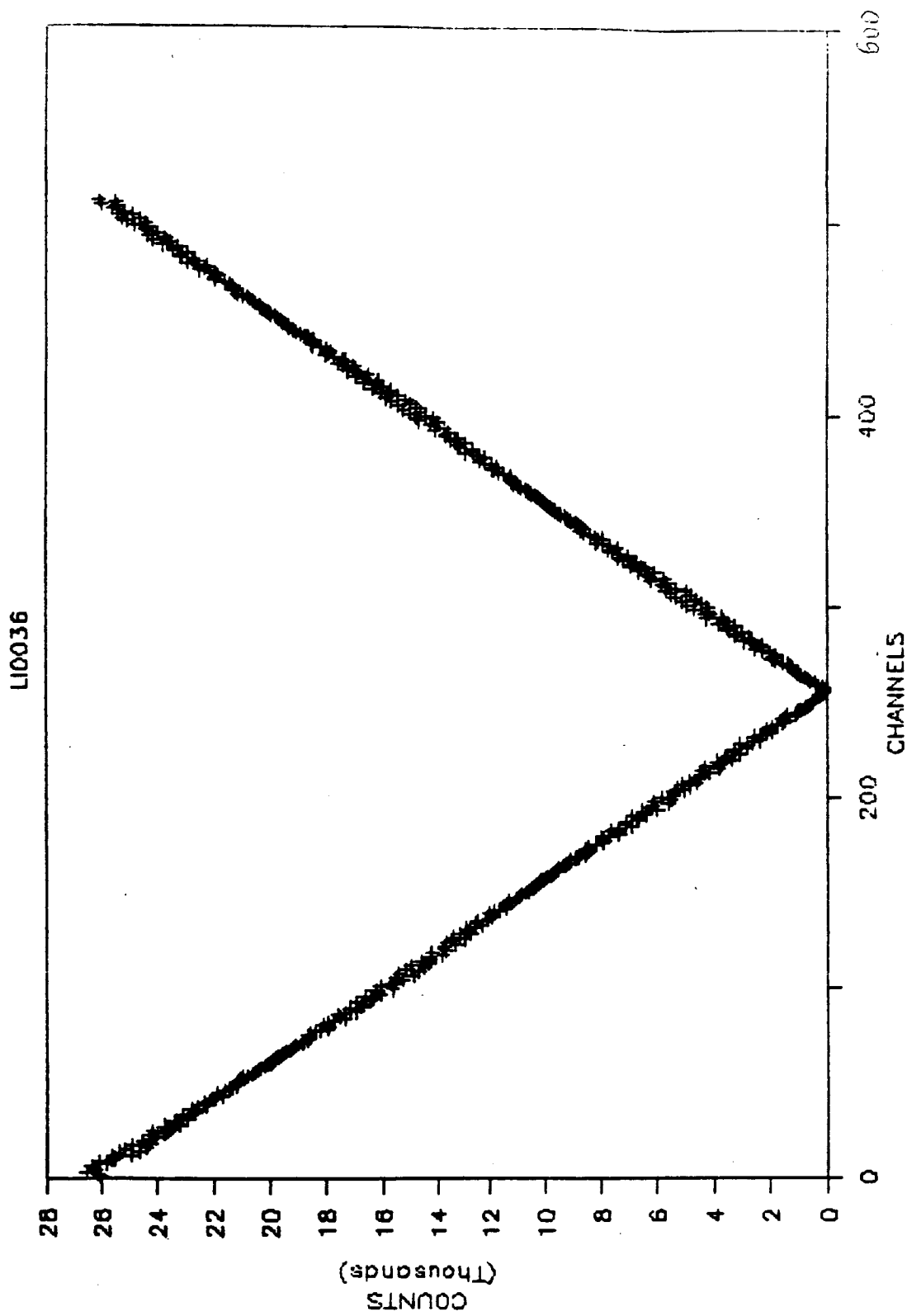
Figure 9. Laser calibration data associated with the spectra of Figures 1,2,11,12.

Iteration	χ^2 by Taylor's	χ^2 by Marquardt
Start	1.126	1.126
1	1.114	1.054
2	1.103	1.010
3	1.094	0.997
4	1.086	0.990
5	1.079	0.987
6	1.072	0.985

Iteration	χ^2 by Taylor's	χ^2 by Marquardt
Start	32.622	32.621
1	18.319	5.833
2	9.957	4.676
3	5.847	1.333
4	1.508	1.240
5	1.483	1.078
6	1.566	1.051
7	1.555	1.023
8	1.541	1.021

Figure 10. Variation of χ^2 for the first few iterations in fitting 6 peaks to the data of Figure 6 by Taylor's or Marquardt procedure. In the upper table, the fit was started relatively close to the minimum of 0.966. In the lower table, the fit was started farther away from the minimum.
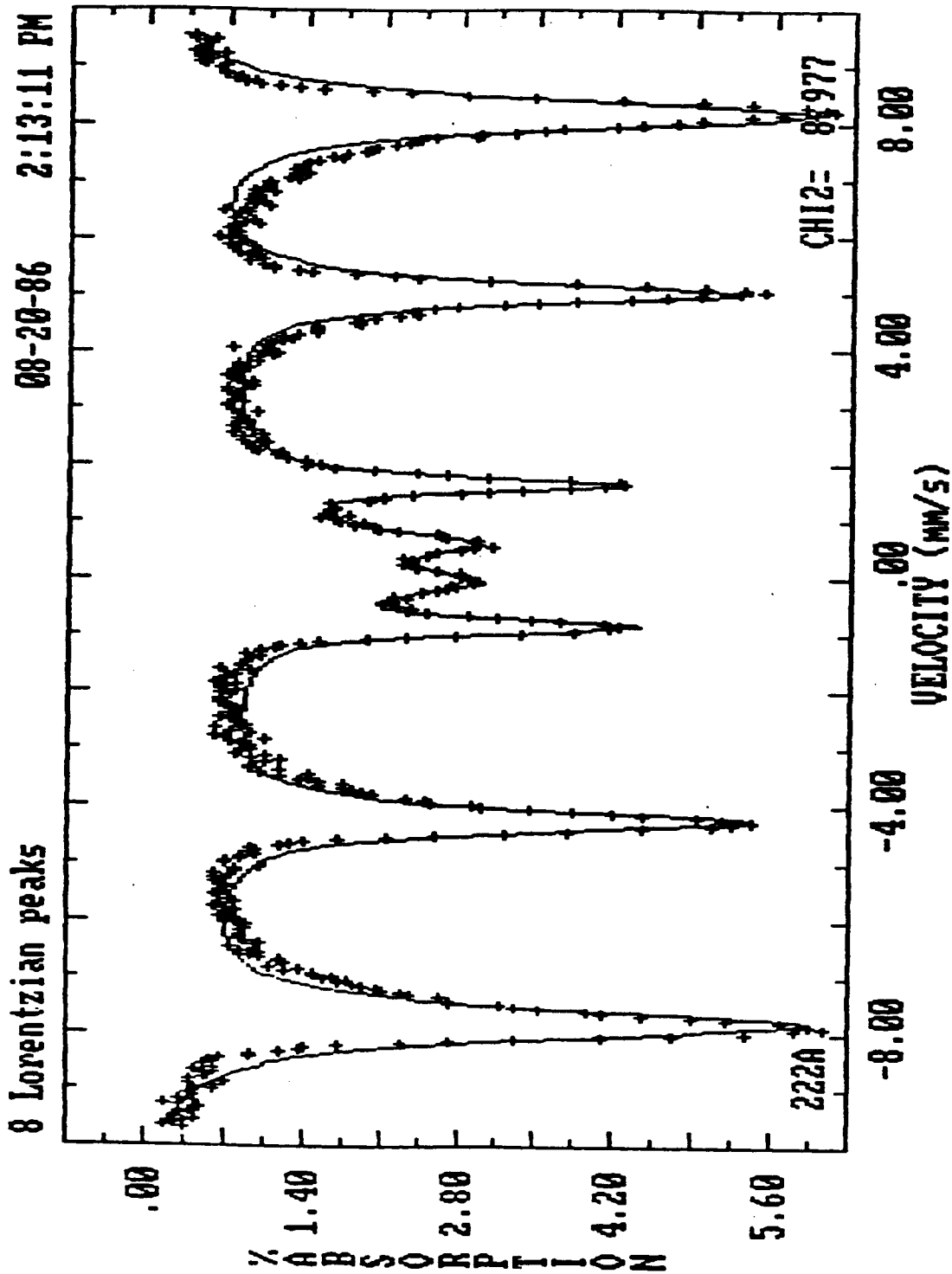
Figure 11. Data of Figure 2 fit to 8 Lorentzian peaks with areas and widths constrained in pairs.
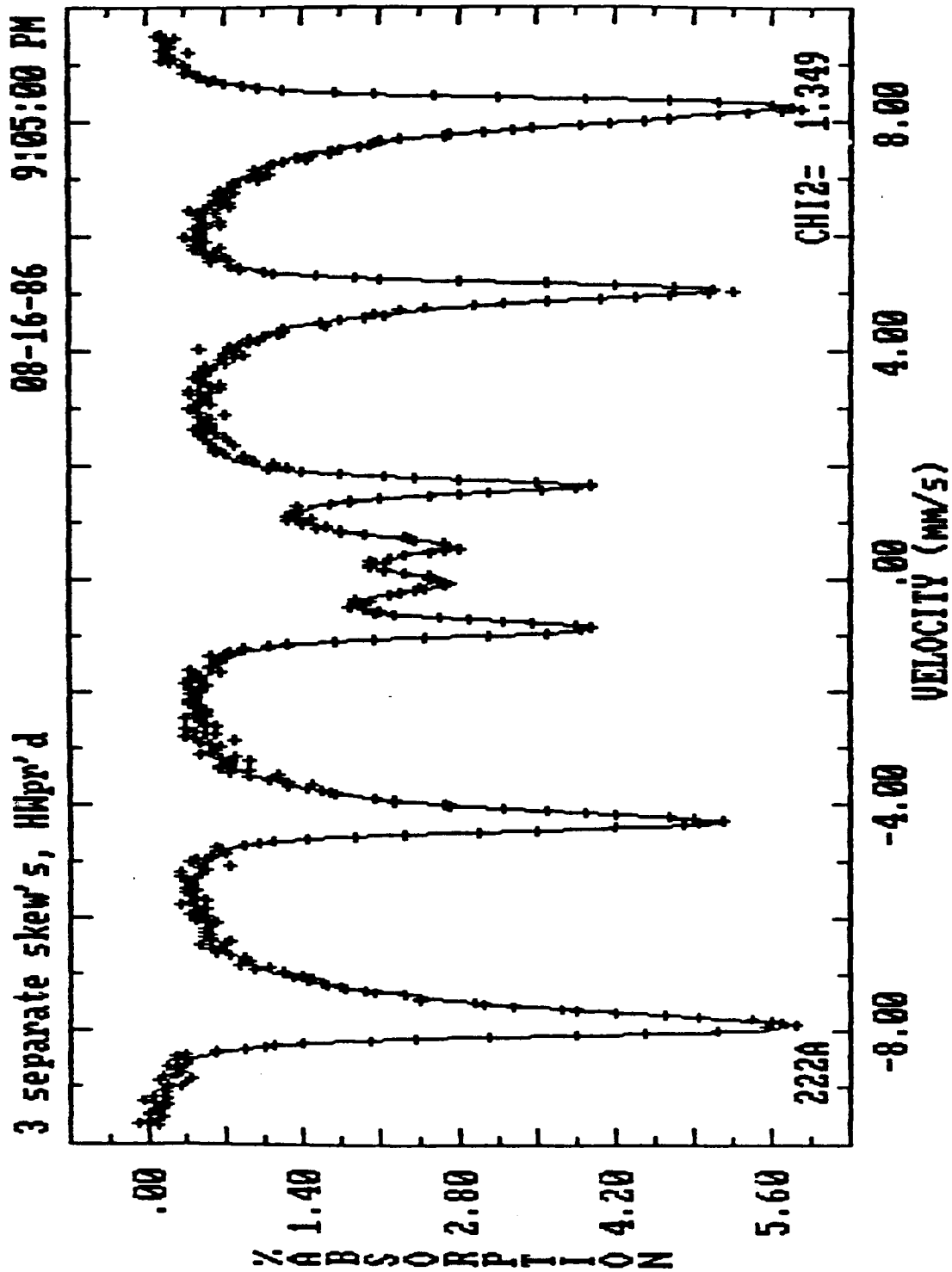
Figure 12. Fit as in Figure 11 with 3 additional "skew" parameters for the 6 magnetic peaks.

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

University of Houston

Vibrational and Rotational Analysis of the Emission

Spectra of the Arc Jet Flow

Prepared by:	Charles Blount, PhD
Academic Rank:	Professor
University and Department:	Texas Christian University
	Department of Physics

NASA/JSC

Directorate:	Engineering
Division:	Advanced Program Office
Branch	Aeroscience
JSC Colleague	Carl D. Scott, Ph.D.
Date	
Contract #:	NGT-44-005-803
	University of Houston

VIBRATIONAL AND ROTATIONAL ANALYSIS OF THE EMISSION

SPECTRA OF THE ARC JET FLOW

Charles E. Blount

Professor

Department of Physics

Texas Christian University

Fort Worth, TX   76129

Applying atomic and molecular physics to the analysis of the radiation emitted from the arc jet flow provides a means for determining the species and excitation temperature of the constituents of the flow.

The object of this investigation has been a rotational and vibrational analysis of the spectra obtained from the radiation emitted in the shock layer and in the free stream of the arc jet flow, specifically, in the shock layer bands of the First Negative Group of ionized molecular nitrogen $(N_2^+ \; B\,^2\Sigma_u^+ \to X\,^2\Sigma_g^+)$ and in the free stream bands of the γ-system of nitric oxide $(NO \; A^2\Sigma_u^+ \to {}^2\Pi_{\frac{1}{2}\frac{3}{2}})$.

INTRODUCTION

The temperature of a gas containing molecules emitting radiation is reflected in the intensity distribution of a vibrational progression and in the intensity distribution of a rotational progression of the electronic transition. Hence, only the identification of the vibrational and rotational levels associated with the observed bands and their relative intensities are necessary to determine the temperature of the gas.

THEORY

The intensity of the radiation of light which accompanies the passage of a molecule from an upper state m to lower state m is determined by the Einstein transition probability for spontaneous emission between the states A_{mm} the number of molecules N_m having the upper state m populated, and the energy of the emitted radiation for the transition $hc\,\nu_{mm}$. The radiation intensity for a gas of molecules is

$$I_{em.}^{m\,m} = N_m\,h\,c\,\nu_{mm}\,A_{mm} \tag{1}$$

The spontaneous emission coefficient can be expressed in terms of the dipole moment for the transition $R^{m\,m}$ by

$$A_{mm} = \frac{64\pi^4\,\nu_{mm}^3}{3h}\,|R^{mm}|^2 \tag{2}$$

It is known that the total dipole moment for an electronic-vibrational-rotational transition in a band system can be expressed in the Born-Oppenheimer approximation by

$$R^{mm} = R_e \, R_{vib}^{v',v''} \, R_{rot}^{J'J''} \qquad (3)$$

where R_e is the dipole moment for the electronic transition, $R_{vib}^{v',v''}$ is the dipole moment for the vibrational transition from an upper vibrational state v' to a lower state v'' and $R_{rot}^{J'J''}$ is the dipole moment for the rotational transition from an upper rotational state J' to a lower state J''.

The expression for the spontaneous emission coefficient in the Born-Oppenheimer approximation for the rotational levels of a molecular system is

$$A_{mm} = 64\pi^4 \, v_{mm}^3 \sum_{M'M''} \frac{|R_{mm}|^2}{2J'+1}$$
$$= 64\pi^4 \, v_{mm}^3 \, |R_e|^2 |R_{vib}^{v',v''}|^2 \sum_{M'M''} \frac{|R_{rot}|^2}{2J'+1} \qquad (4)$$

where the sum over M' and M'' accounts for all the equivalent spacial orientations of J' and J'' and $2J'+1$ is the statistical weight of the rotational level J'. It will also be noted that the square of the dipole moment for a vibrational transition $|R_{rot}^{v',v''}|^2$ is usually designated by $q_{v',v''}$ and referred to as either the Franck-Condon factor or the square of the overlap integral [1,2]. The sum $\sum_{M'M''} |R_{rot}^{J'J''}|^2$ is

generally denoted by $S^{J'J''}$, the line strength, and referred to as the Höln-London factor [1,2]*.

The intensity of the emitted radiation in terms of the dipole moments for the transition is given by

$$I_{em}^{nm} = \frac{64\pi^4}{3} c N_m \nu_{nm}^4 |Re|^2 |R_{vib}^{nm}|^2 \sum_{M'M''} |R_{rot}^{J'J''}|^2 / 2J' + 1 \quad (5)$$
$$= \frac{64\pi^4}{3} c N_m \nu_{nm}^4 |Re|^2 q_{v'v''} \frac{S^{J'J''}}{2J'+1}$$

If the summation is limited to the transition probabilities over , ie a single electronic-vibrational band, one obtains

$$I_{em}^{nm} = \frac{64\pi^4}{3} c N_m \nu_{nm}^4 |Re|^2 q_{v'v''} \quad (6)$$

_____

*Summation Rules

The sum of the line strength of all transitions from or to a given rotational level is equal to the statistical weight.

The sum of the squares of the overlap integrals summed over all values of the supper or of the lower state is equal to one.

Calling band strength the emission intensity divided by ν^4, it follows from the vibration sum rule that the sum of the band strengths of all bands with the same upper vibrational state are proportional to the number of molecules in the upper state, ie

$$\sum_{\nu''} \frac{I_{em}^{\nu'\nu''}}{\nu^4} \sim N_{\nu'}$$

(7)

One can readily see that the above derivation is valid only if the electronic transition dipole moment R_e is a constant for all vibrational transitions having band strengths that give an appreciable contribution to the sum. For molecular systems satisfying this condition the sum rule may be used to determine the temperature of the gas emitting the band system. In thermal equilibrium the population $N_{\nu'}$ of the initial state is proportional to $e^{-\frac{G(\nu')hc}{kT}}$. We obtain from (7)

or

$$\sum_{\nu''} \frac{I_{em}^{\nu'\nu''}}{\nu^4} \sim e^{-\frac{G(\nu')hc}{kT}}$$

(8)

$$\ln\left(\sum_{\nu''} \frac{I_{em}^{\nu'\nu''}}{\nu^4}\right) = C - G(\nu')hc/kT$$

Therefore by plotting the logarithms of the sum of the band strengths against the vibrational term values a straight line is obtained whose slope is $\frac{hc}{kT}$. However, if the intensities of sufficient number of bands cannot be measured a determination of the vibrational temperature can be obtained if the Franck-Condon factors have been calculated for the measured bands. Dividing the band strength of each band by its Franck-Condon factor gives again a quantity that is

proportional to the number of molecules in the initial state and may be plotted in the same way as the above.

The intensity of the lines in the rotational branch of an electronic-vibration band is given by

$$I_{em}^{J'J''} = N_{J'} \, \nu^4 \, \frac{S^{J'J''}}{2J'+1} \tag{9}$$

one can readily see from (5) that the above expression is valid only if the electronic-vibration dipole moment is constant for all the rotational transitions that give an appreciable contribution to the branch.

In thermal equilibrium the population $N_{J'}$ of the initial state is proportional to $(2J'+1) \, e^{-F(J') \, hc/kT}$. We obtain from (9)

$$\frac{I_{em}^{J'J''}}{\nu^4 S^{J'J'}} \sim e^{-F(J')hc/kT} \tag{10}$$

Please note that by assuming ν to be constant over the entire branch an error of less than one percent is introduced and the above expression reduces to

or
$$\frac{I_{em}^{J'J''}}{S^{J'J''}} \sim e^{-F(J')hc/kT}$$

$$\ln\left(\frac{I_{em}^{J'J''}}{S^{J'J''}}\right) = C - F(J') \frac{hc}{kT} \tag{11}$$

In the special case of $^1\Sigma \rightarrow {}^1\Sigma$ transitions for which there is only a P and R branch the intensity relations for emission are given by

$$I_{em} \sim (J' + J'' + 1)\, e^{-B'J'(J'+1)\frac{hc}{kT}}$$

or   (12)

$$\ln\left(\frac{I_{em}}{J'+J''+1}\right) = C - B'J'(J'+1)\frac{hc}{kT}$$

(13)

Therefore, by plotting $\ln\left(\frac{I_{em}}{J'+J''+1}\right)$ against $J'(J'+1)$ a straight line is obtained whose slope is $-\frac{B'hc}{kT}$. Thus if the line intensities have been measured and the rotational constant is known the temperature of the gas may be determined.

DISCUSSION AND RESULTS

From the spectra obtained in the shock layer of the arc jet flow, the R branch of the (0,1) band of the First Negative Group of ionized molecular nitrogen N_2^+ $B^2\Sigma_u^+ \rightarrow X^2\Sigma_g^+$ has been selected for the determination of an effective rotational temperature. Although both the upper and lower states are doublets, the doublets are not resolved in the R branch of either the (0,0) or (0,1) band, figure 1. Therefore the tempterature can be determined using the special case $^1\Sigma \rightarrow {}^1\Sigma$ given by equation (13) using the K values for the rotational band assignments rather than the J values. Please note that if rotational branchs of bands other than (0,0) or (0,1) are selected for the determination of

the temperature the doublet character of the spectra may be resolved and equation 13 is no longer valid for temperature determinations, figure 2.

In figure 1, note the alternation of the intensity between even K, labeled, and odd K. This alternation of intensity is due to the nuclear spin of the nitrogen nuclei which will result in twice as many nolecules with even spin as with odd resulting in the observed intensity alternation in the ratio 2:1. For a temperature determination one may either use the even K intensities or the odd K intensities but not a mixture.

For the R branch of the (0,1) band values of Te = 25461.46, We = 2419.84, wexe = 23.189 and weye = -0.537 for the electronic and vibrational terms of the $B^2\Sigma_u^+$

state; and Te = 0, we = 2207.00, wexe = 16.1, and $\omega_e y_e$ = -0.04 for the electronic and vibrational terms of the $X^2\Sigma_g^+$ state were obtained from Huber and Herzberg (4).

The rotational values were obtained from the data reported by K. A. Dick, et al (5) using the combination relations,

and

$$R(J-1) - P(J+1) = (4B_{v''} - 6D_{v''})(J+\tfrac{1}{2}) - 8D_{v''}(J+\tfrac{1}{2})^3$$

$$R(J) - P(J) = (4B_{v'} - 6D_{v'})(J+\tfrac{1}{2}) - 8D_{v'}(J+\tfrac{1}{2})^3$$

between the lines of the P and R branch for the evaluation of rotational constants for bands without a Q branch. For the (0,1) band values of

$B_v' = 2.0751$, $D_v' = 6.54 \times 10^{-6}$, $B_v'' = 1.9031$, and $D_v'' = 5.8 \times 10^{-6}$ were obtained. The energy in cm^{-1} for the R branch of the (0,0) and (0,1) bands of the $B^2\Sigma_u^+ \to X^2\Sigma_g^+$ transition of N_2^+ can be fit to the measured values of Dick, etal (5) to within $1 cm^{-1}$ using

$$\nu_R = 25,570.21 + 4.301 K + 0.1512 K^2$$

for the (0,0) band and

$$\nu_R = 23,395.64 + 4.296 K + 0.1718 K^2$$

for the (0,1) band.

In table 1 values used for the determination of the rotational temperature in the shock layer of the arc jet flow are given. The relative intensity of the lines are given in arbitrary units and were obtained from densitometer tracings of plate #14 using the Joyce-Lobel Microdensitometer in the Spectroscopy Laboratory of the Department of Physics at Texas Christian University, Fort Worth, Texas.

Table I

K'	K''	I	$I/(K'+K''+1)$	$\ln\left(\dfrac{I}{K'+K''+1}\right)$	$K'(K'+1)$
1	0	193	96.5	4.57	2
3	2	224	37.3	3.62	12
5	4	267	26.7	3.28	30
7	6	292	20.9	3.04	56
9	8	309	17.2	2.84	90
11	10	309	14.1	2.65	132
13	12	261	10.0	2.30	182
15	14	295	9.83	2.29	240
19	18	300	7.90	2.07	380
21	20	290	6.90	1.93	462
23	22	296	6.43	1.86	552
2	1	121	30.3	3.41	6
4	3	206	25.8	3.25	20
6	5	192	16.0	2.77	42
8	7	237	14.8	2.69	72
10	9	224	11.2	2.42	110
12	11	245	10.2	2.32	156
14	13	208	7.43	2.01	210
20	19	271	6.78	1.91	420
22	21	255	5.80	1.76	506
24	23	237	4.94	1.60	600

A graph of $\ln\left(\dfrac{I}{K'+K''+1}\right)$ vs $K'(K'+1)$ from values given in table 1 for even K'' and for odd K'' is shown in figure 3. Note, that for both even and odd K'' the points for the lower values of K'' are found to be above the line obtained for the higher values of K''. From the energy terms of the P and R branches

$$\nu_P = \nu_{01} - (B_{v'} + B_{v''})K + (B_{v'} - B_{v''})K^2$$

and

$$\nu_R = \nu_{01} + 2B_{v'} + (3B_{v'} - B_{v''})K + (B_{v'} - B_{v''})K^2$$

it is seen that at large K the K^2 term is dominant. This will place members of the P branch having high K values in the region of the R

branch of low K values. Thus the observed intensity in the R branch for low K values is both P and R. This will not present any problems in the determination of the temperature since at lower temperature the intensity of the P branch will decrease for high K values and the intensity observed in the R branch at low K values will be due to R only.

The slope of the line was found to be -0.00148 for even K and . -0.00165 for odd K. Using the value $B_v' = 2.0751$ a temperature of 2020 ^oK was obtained for even K and 1810 ^oK for odd K.

CONCLUSIONS

The reults presented above indicate that the temperature of the arc jet flow can be determined from the measurements and conditions presented. I regret that the time was not sufficient for the determination of the temperature in the free stream of the arc jet flow from the vibrational analysis of the system of NO. However, the partial analysis has indicated that the intensities of bands having wavelengths less than 2600 are antenuated. Also, bands are observed in the second order that are not observed in the first order. The blaze of the grating is suspect indicating a blaze for wavelengths much greater than the region of interest. Also, results on the system of NO- by N. E. Kuz' menko, etal (6) report a dependence of the electrnoic dipole moment on the vibrations of the molecule. Values for the

electronic dipole moment for each of the bands observed in the free can be obtained from this article and with the Frank-Condon factor for each band the temperature can be obtained.

ACKNOWLEDGEMENTS

References

1. G. Herzberg, <u>Molecular Spectra and Molecular Structure I. Spectra of Diatomic Molecules.</u> D. Van Nostrand, New York, 1950.

2. J. I. Steinfeld, Molecules and Radiation: An Introduction to Modern Spectroscopy, Harper and Row, New York, 1974.

3. D. V. Skobel'tsyn, <u>Electronic and Vibrational Spectra of Molecules, Consultants Bureau</u>, New York, 1968.

4. K. P. Huber and G. Herzberg, <u>Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules</u>, D. Van Nostrand, New York, 1979.

5. K. A. Dick, W. Benesch, H. M. Crosswhite, S. G. Tilford, R. A. Gottscho, and R. W. Fields, J. Mol. Spectros. 69, 95, 1978.

6. N. E. Kuz'menko, L. A. Kuzetsova, A. P. Monyakin, and Yu Ya Kuzyakov, J. Quant. Spectrosc. Radiat. Transfer, 24, 219, 1980.
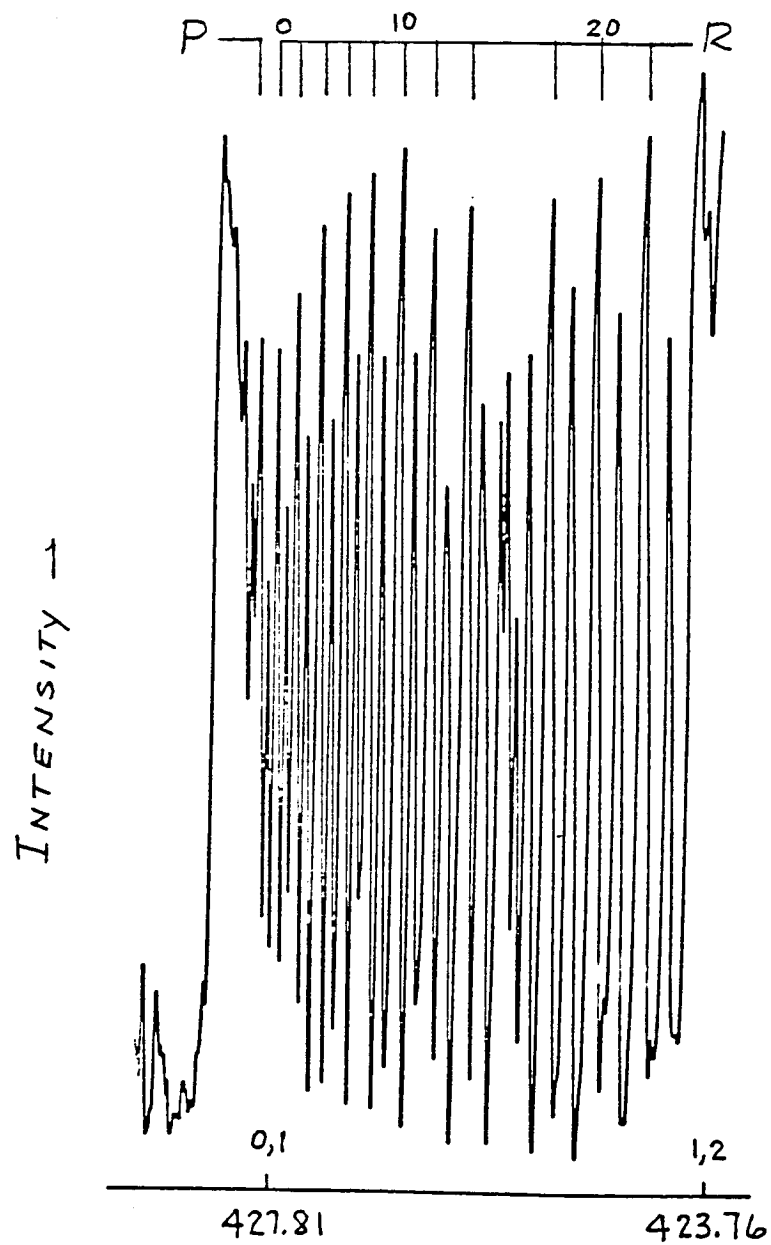
Figure 1. Densitometer tracing of the (0,1) band in the γ system of NO with even values of K for the R branch indicated.
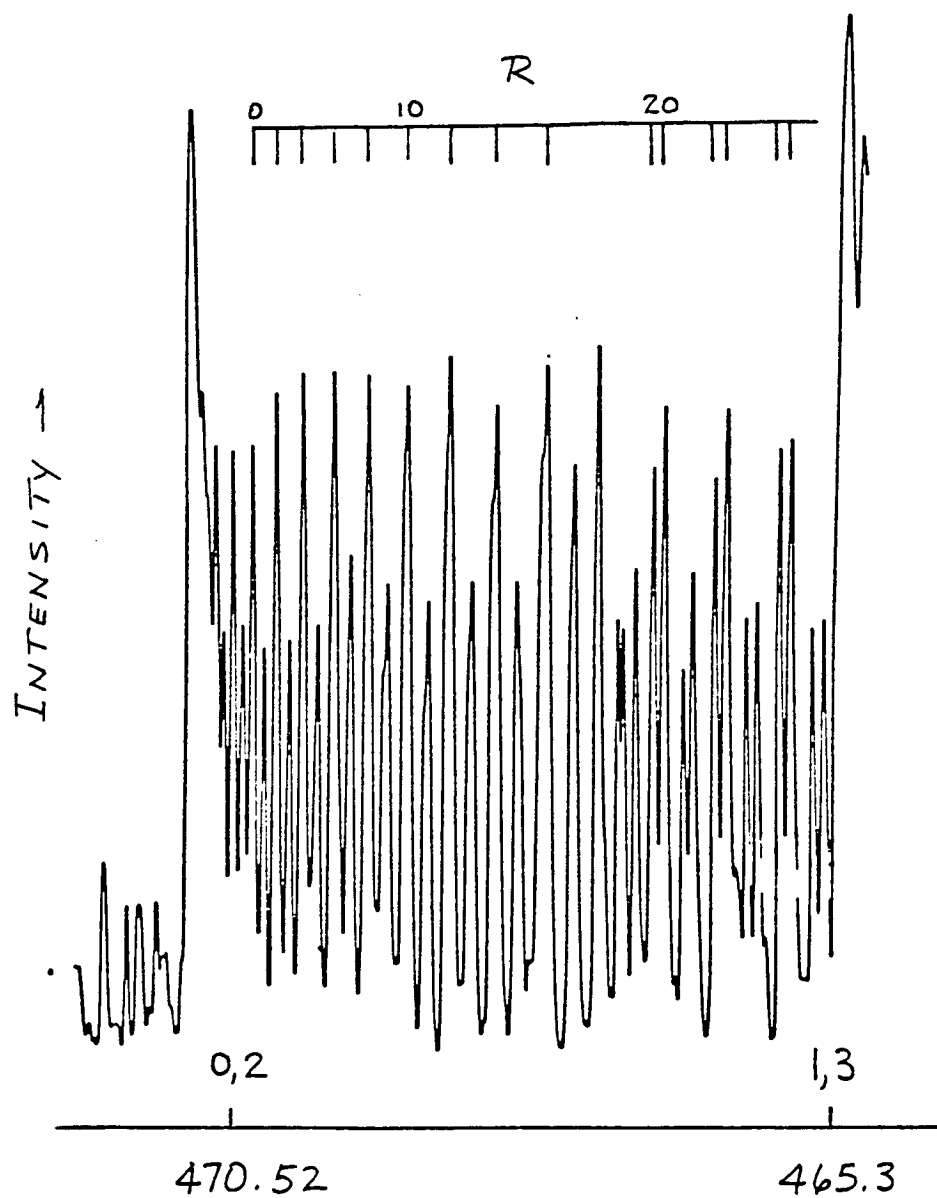
Figure 2.   Densitometer tracing of the (0,2) band of
            the γ system of NO with even values of
             K for the R branch indicated and the
            doublet structure for large values of K
            resolved.
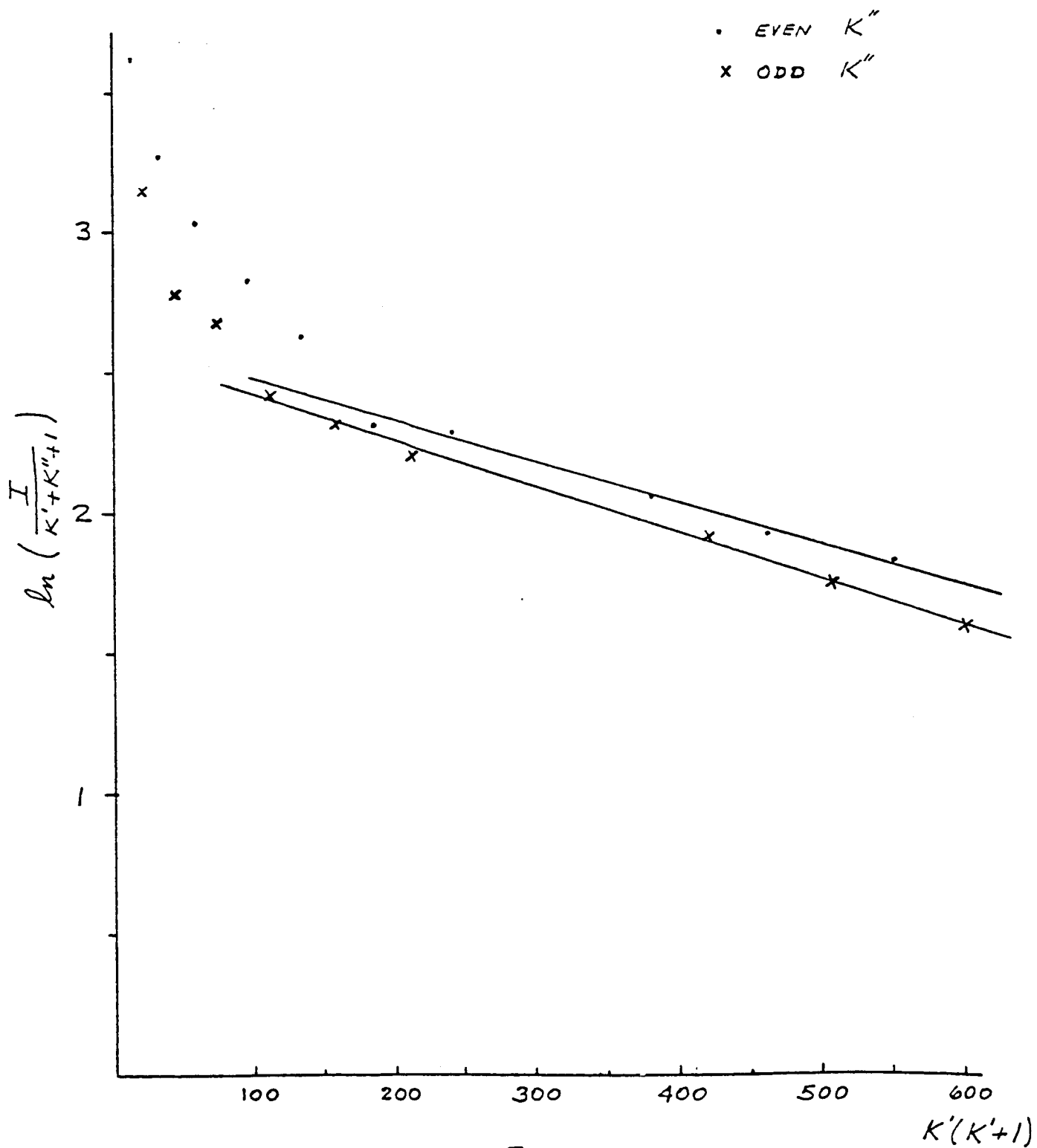
Figure 3. Plot of the $\ln\left(\frac{I}{K'+K''+1}\right)$ vs $K'(K'+1)$ for even and odd K of the R branch of the (0,1) band.

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

University of Houston

Distributed Phased Array Architecture Study

Prepared by: Brian Bourgeois, Ph. D.

Academic Rank: Assistant Professor

University & Department: University of Houston-Downtown

Applied Mathematical Sciences

NASA/JSC

Directorate: Engineering

Division: Tracking and Communications

Branch: Electromagnetic Systems

JSC Colleague: George D. Arndt, Ph.D.

Date: August 8, 1986

Contract #: NGT44-005-803

DISTRIBUTED PHASED ARRAY ARCHITECTURE STUDY

Brian Bourgeois, Ph.D
Assistant Professor
Department of Applied Mathematical Sciences
University of Houston-Downtown
Houston, TX   77002

The hardware tolerances needed to successfully operate distributed phased array antennas in a space environment are not clearly defined at this time. Variations in amplifiers and phase shifters can cause degraded antenna performance, depending also on the environmental conditions and antenna array architecture.

The implementation of distributed phased array hardware has been studied with the aid of the DISTAR computer program as a simulation tool. The principal task of this simulation is to provide guidance in hardware selection. Both hard and soft failures of the amplifiers in the T/R modules are modeled. Hard failures are catastrophic – no power is transmitted to the antenna elements. Non-catastrophic or soft failures are modeled as a modified Gaussian distribution. The resulting amplitude characteristics then determine the array excitation coefficients. The phase characteristics take on a uniform distribution.

Pattern characteristics such as antenna gain, half-power beamwidth, mainbeam phase errors, sidelobe levels, and beam pointing errors have been studied as functions of amplifier and phase shifter variations. General specifications for amplifier and phase shifter tolerances in various architecture configurations for C-band and S-band have been determined.

NASA Collegue: G. Dickey Arndt   EE3   X2128

INTRODUCTION

The distributed architecture concept in phased array antennas incorporates transmit/receive (T/R) modules at or near the elemental radiators of the array. The most important components of the T/R modules are the high power amplifier (HPA) and the low noise amplifier (LNA) . Major advantages of this approach include system reliability, improved system noise figure, mechanical deformation and motion compensation, and achievement of high total radiated power with solid state devices.

The most generic distributed array has an amplifier (or T/R module) at each radiating element. Due to limitations of cost or practicality, the array architecture may require reduction, so that one module may drive several elemental radiators. An important problem is to optimize antenna performance subject to the constraint of architecture reduction. Further constraints include the use of real rather than ideal electrical components, which are subject to both random and systematic errors.

To address this problem, a computer program named DISTAR has been created by PSL (Physical Sciences

Laboratory, New Mexico) and developed by NASA/JSC. The program inputs antenna array characteristics along with type and extent of amplifier performance failure and outputs the normalized antenna gain pattern in graphical and/or tabular form. Both hard and soft failures of the amplifiers in the T/R modules are modeled. Hard failures are catastrophic - no power is transmitted to the antenna elements. Soft failures are random perturbations of amplitude and phase from the ideal specifications.

The paper gives a brief description of the program DISTAR, followed by an analysis of the method used to construct the pattern. The final section discusses an application of the program to determine specifications for hardware tolerances for three distributed arrays, one at C-band and two at S-band.

PROGRAM DESCRIPTION

This section briefly describes the capability of
the program DISTAR in terms of input and output. The
array is rectangular. It may be divided, both physically
and electronically, into various subarrays: panels,
subgroups, co-phased elements, and co-amplified
elements. The dimensions of these subarrays are all
determined by the user. It may be useful to refer to
Figure 1, which sketches a 12 x 6 element array with 6
panels and 3 x 2 element subgroups. The co-amplified
groups are the panel rows.

Each panel is excited in amplitude and phase by
user-specified amounts. A panel must contain an integral
number of subgroups and co-phased groups. Each subgroup
is physically separated from its neighbors by a uniform
amount in x and y. Each element in a co-phased group is
given an identical phase shift. Co-amplified elements
are all driven by the same T/R module. The user
specifies the spacing in x and y between elements and
between subgroups, the frequency of the antenna, the
element taper, the element pattern, the steering angle,
and display mode(s) (2D graphs, 3D graphs, table).

Information about type and degree of hardware
failure is input via program flags. If the user requests
soft failures of the T/R modules, the program prompts
for mean power, standard deviation in power, and range
of phase distribution. (See next section for more
detail.) If the user requests hard failures, the program
prompts for whether the modules should be turned off
randomly or systematically. If systematically, the user
supplies the number turned off. If randomly, the user
chooses whether to supply the number or have it also
selected randomly.

GEOMETRY

12 X 6 ELEMENT ARRAY   3 X 2 ELEMENT SUBGROUPS

FIGURE 1

3-7
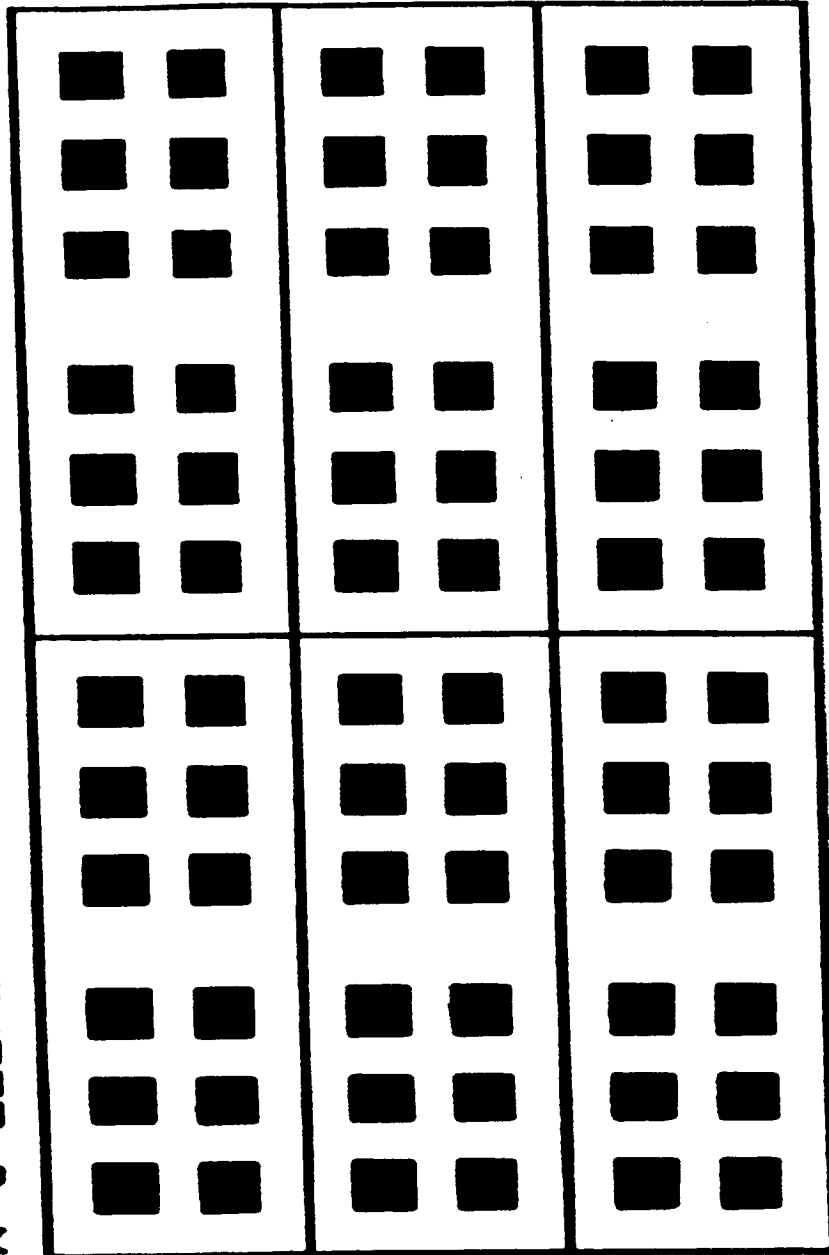
THEORY

In this section, the equations used by the program to calculate the GAIN matrix are detailed. A brute-force method is used to sum the contributions of all the antenna elements to the field in a given direction. The GAIN matrix is calculated exactly once in the program and is subsequently used to display the information in the various forms requested by the user. For the convenience of the interested reader, the notation used in this section is identical to that used in the program.

For a given THETA and PHI, the linear complex array directivity AF2 is calculated in subroutine ARRAY as a sum over the contributions from the panels (see Section 1)

$$AF2 = \sum_{panels} A1 \star SUBEF \star EXP(iA2) \; ,$$

where

    A1 = panel amplitude excitation coefficient

    A2 = panel phase excitation coefficient

    SUBEF = panel complex electric field

The array factor is given by

$$AF = |AF2|^2 \star F / (MEL \star NEL \star POUT \star XNORM) \; ,$$

where

MEL = the number of elements per panel in the
      x-direction

NEL = the number of elements per panel in the
      y-direction

$$\text{POUT} = \sum_{\text{panels}} (A1)2$$

$$\text{XNORM} = \sum_{\text{all elts}} (\text{ELWT})2 \,/\, \#\text{elts}$$

ELWT = matrix containing the weights from the
       element taper

$$F = \begin{cases} (1/16)[\ 1-\cos(\text{PI}-\text{THETA})\ ]4 & \text{if IELP = 1} \\ \\ 1 & \text{if IELP = 0} \end{cases}$$

IELP = the element pattern flag

Then,

PHAS(THETA,PHI) = the complex argument in degrees
                  of AF2

and

GAIN(THETA,PHI) = 10LOG_{10} (AF) = AF expressed in
                  decibals.

The panel electric field SUBEF is calculated in subroutine SUBARY as follows:

$$\text{SUBEF} = \sum_{\substack{\text{elts in} \\ \text{panel}}} z \ w \ x \ P \ A \ ,$$

where

$$z = \text{ELZAP} = \begin{cases} 0 \text{ if element is zapped} \\ \quad \text{(catastrophic failure)} \\ \\ 1 \text{ if element is not zapped} \end{cases}$$

$w = \text{ELWT} = $ weight from the element taper

$x = \text{EXPHAS} = $ relative phase shift of excitation to

steer the beam to THETA0,PHI0.

x is a complex number of modulus one.

THETO,PHIO is the pointing angle.

$P = \text{PHASE} = $ phase at current look angle. P is a

complex number of modulus one

$A = \text{AMPWT} = $ amplitude weight which models soft

failures, as described below.

The amplitude weight $A = \text{AMPWT}$ is calculated in subroutine AMPLWT as follows:

$$A = (a/u)^{1/2} \ \text{EXP(PHS)} \ ,$$

where

$$a = u + (-2*VAR*\ln X1)1/2 \cos(2*PI*X2)$$

PHS = -j*DELTA*(1-2*X3) = uniform distribution

between -DELTA and DELTA

u = mean of the distribution

(user-supplied = AMEAN)

VAR = variance = SG*SG = square of standard

deviation SG

(SG is user-supplied)

DELTA = range of phase distribution (user-supplied)

X1,X2,X3 are randomly generated real numbers

between 0 and 1.

ANTENNA TESTS

The program DISTAR described above was used to test three antennas for NASA, two at S-Band and one at C-Band. The problem was to determine the hardware tolerances necessary to operate these antennas in a space environment. With this model, this means to determine to what degree the amplifiers in the T/R modules can fail and still maintain an adequate antenna performance.

Two straightforward criteria were established to determine the hardware tolerances. First and foremost, the power at the maximum of the degraded beam should be within three decibals of the power of the maximum of the ideal beam. In other words, a falloff in power of more than fifty percent is not tolerated. Second, sidelobes of the degraded beam should not rise to within ten decibals of the mainlobe in the degraded beam.

Both hard and soft failures of the T/R modules were tested. Soft failures included both amplitude and phase errors. Different steering angles were employed. Warping of the panels was not included in the study. Principal plane cuts were obtained for all tests.

18 × 12 element   C-Band

The frequency of this microstrip panel was 5.3 GHz. The spacing of the elements was 4.0 centimeters in the x-direction and 3.5 centimeters in the y-direction. Twelve T/R modules were employed, each controlling the eighteen elements in a row of the array. For the random fluctuations, the mean power was 10 decibals, with standard deviation 1 decibal and phase range distribution 10 degrees. The tests were run for two steering angles, i.e., broadside and $\Theta = 20$, $\Phi = 90$. Θ is the polar angle from the z- axis, and Φ is the azimuthal angle measured counterclockwise in the plane of the antenna from the x-axis. The conclusions for hardware tolerances were nearly identical for the two steering angles.

The conclusions are as follows:

1)   Soft failures (random fluctuations in both amplitude and phase) have virtually no effect on the radiation pattern. One reason for this is that the fluctuations were small, the standard deviation of the amplitude variation being 10 percent of the mean, and the phase discrepencies being within 10 degrees.

2) The maximum acceptable level of hard failures is two. Beyond that, there is a high degree of probability that one or both of the above criteria will not be met. The degradation of the pattern is greatest when the failures are concentrated at the center of the antenna. With two hard failures, there is a very small probability that the sidelobes in the elevation plane will rise to within 10 decibals of the mainlobe.

2 x 4 element S-Band

Microstrip panels at two different frequencies were tested at S-Band. The frequencies were 2.1064 GHz and 2.2875 GHz. Since the results for the two frequencies are almost identical, only those of the former antenna will be reported here.

The spacing of the elements was 0.47 λ in the x-direction and 0.56 λ in the y-direction, where the wavelength λ equals 14.242 centimeters. Each array element was controlled by an independent T/R module. For the random fluctuations, the mean power was 7 watts, with standard deviation 0.5 watts and phase range

distribution 25 degrees. Degraded patterns were desired for three different steerings: 1) broadside; 2) Θ = 90 degrees, ϕ = 0 degrees; 3) Θ = 45 degrees, ϕ = 90 degrees.

It was discovered that the antenna could not be steered to the directions 2) and 3) above. The maximum angle in Θ to which the beam can be steered is about 10 degrees. The probable cause for this phenomenon is a combination of two factors:

a) the small number of elements;

b) the element pattern $F = \{ (1/2)[1-\cos(\Pi - \Theta)] \}^4$.

The array factor produced by a) is not strong enough offset the contribution of b) at small values of   .
The ratio of the element pattern for Θ = 0 degrees to that for Θ = 90 degrees is 16.

The conclusions for the broadside tests are as follows:

1) Soft failures have a negligible effect (less than 1 percent) on the maximum power levels due to the small standard deviation of 0.5 watts compared to the mean of

7 watts. However, they appear in some tests to contribute to a small (less than 1 degree) drift of the mainlobe and, when combined with hard failures, to undesirably high sidelobe levels.

2) The maximum acceptable level of hard failures is two. With three hard failures, the average loss in decibals at the maximum is greater than 4. With two hard failures, the average loss in decibals is between 2.5 and 2.6 , with one pattern measured at 2.96 . With soft failures, there is about a 20 percent chance that a sidelobe could rise to within 10 decibals, even within 6 decibals.

Graphical displays of the results are given in Figures 2-6. Since the gain shown is normalized, however, one must examine tabular output to determine absolute power levels.

NORMALIZED ANTENNA GAIN VS. THETA
PHI = 0.

NORMALIZED ANTENNA GAIN VS. THETA
PHI = 90.

IDEAL

2 X 4 ELEMENT   S-BAND

FIGURE 2

NORMALIZED ANTENNA GAIN VS. THETA
PHI = 0.

NORMALIZED ANTENNA GAIN VS. THETA
PHI = 90.
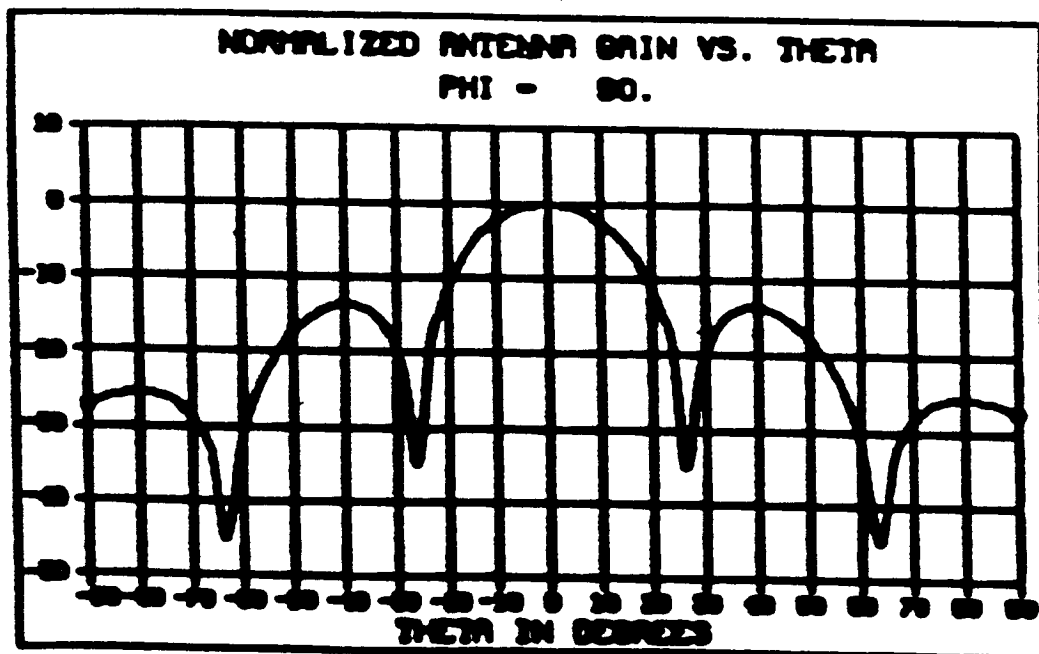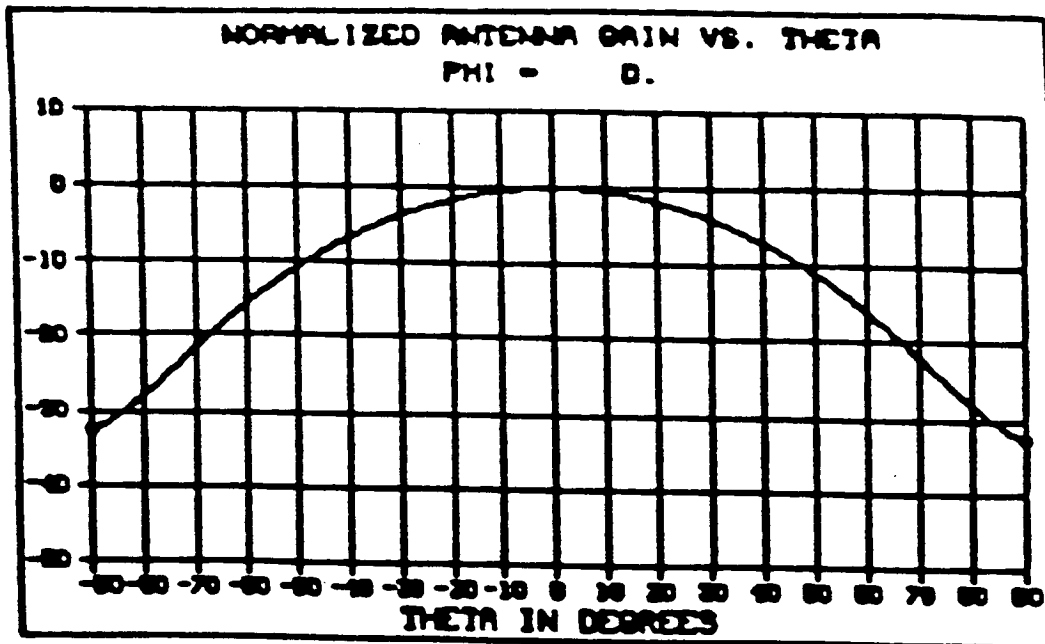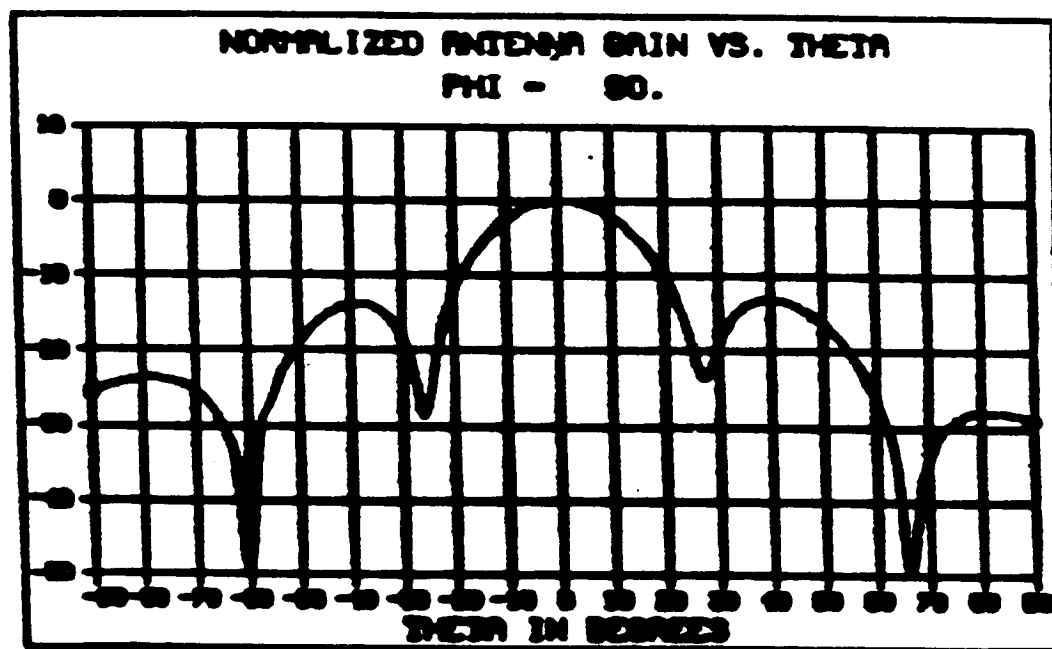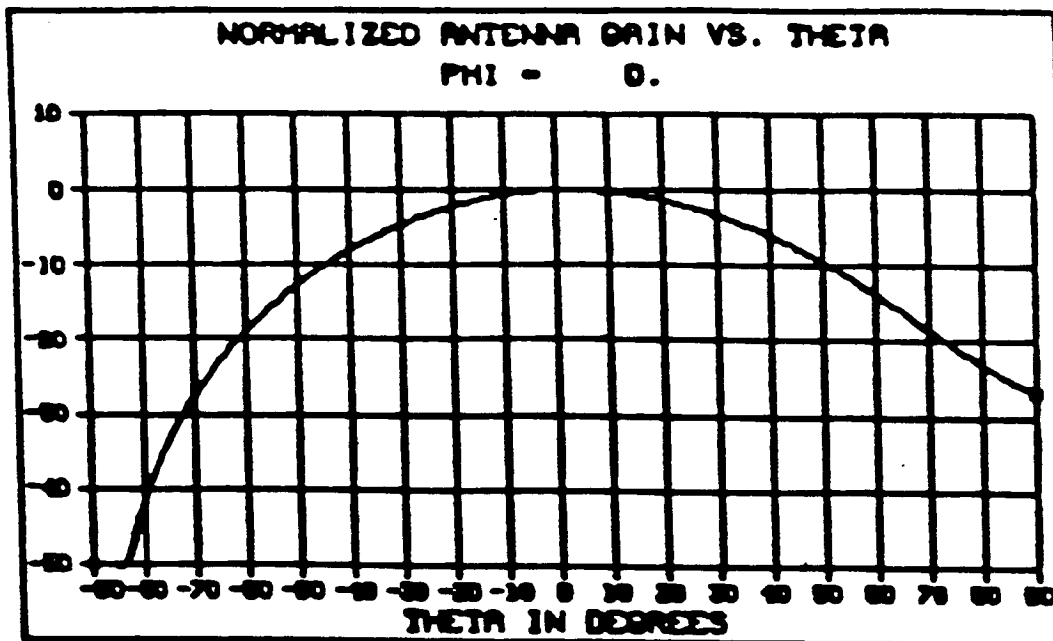
SOFT FAILURES ONLY

2 X 4 ELEMENT   S-BAND

FIGURE 3

NORMALIZED ANTENNA GAIN VS. THETA
PHI = 0.

THETA IN DEGREES

NORMALIZED ANTENNA GAIN VS. THETA
PHI = 90.

THETA IN DEGREES
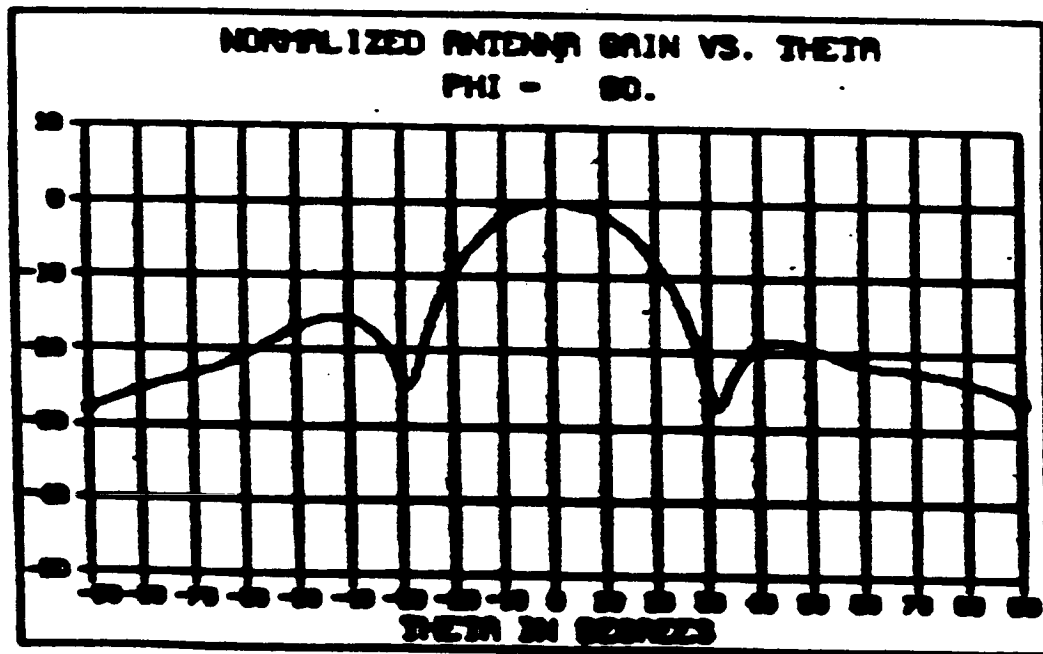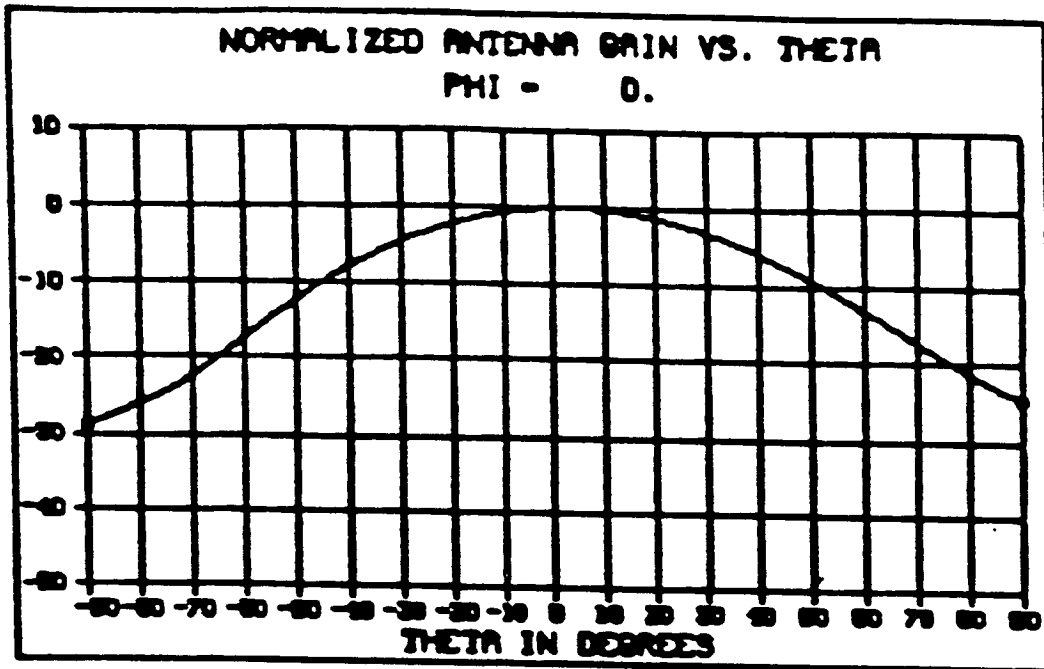
1 HARD FAILURE, SOFT FAILURES

2 X 4 ELEMENT   S-BAND
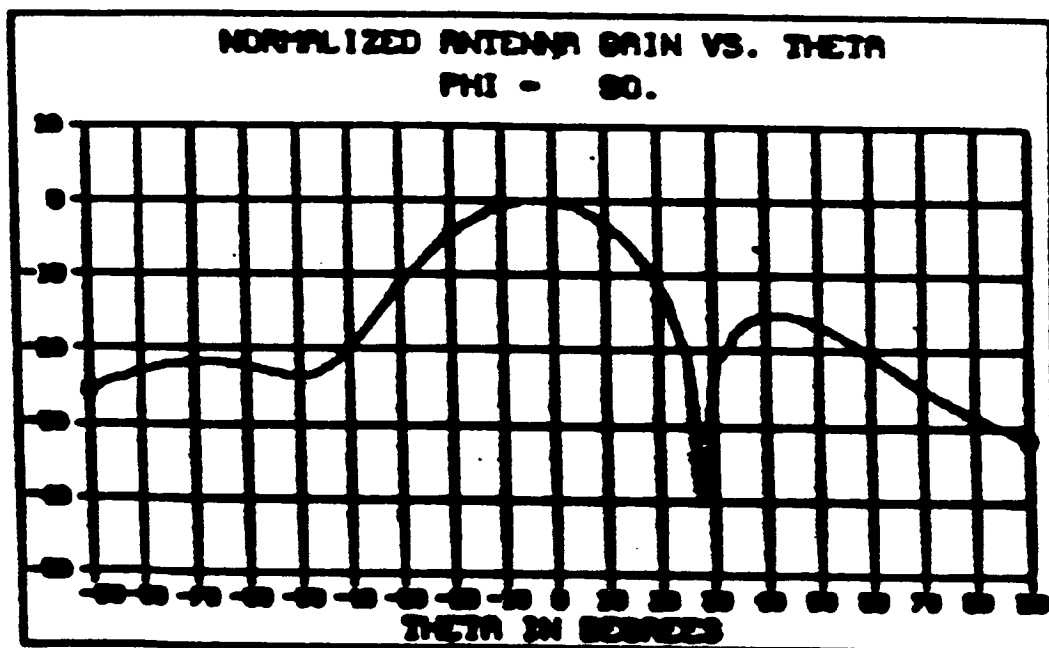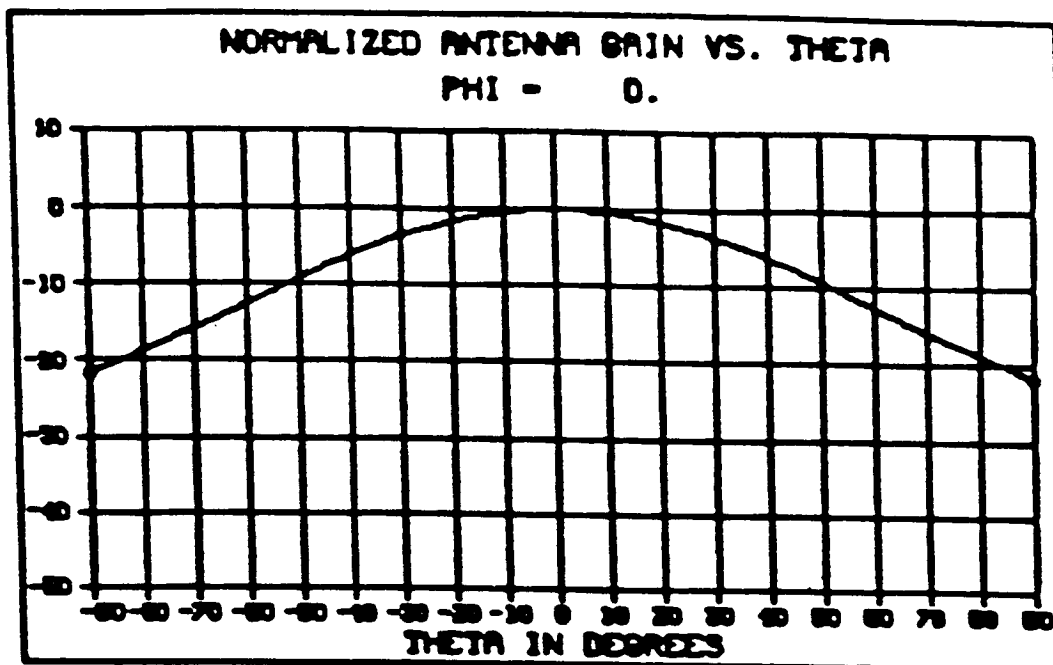
FIGURE 4

NORMALIZED ANTENNA GAIN VS. THETA
PHI = 0.

NORMALIZED ANTENNA GAIN VS. THETA
PHI = 90.

2 HARD FAILURES, SOFT FAILURES

2 X 4 ELEMENT   S-BAND
FIGURE 5

NORMALIZED ANTENNA GAIN VS. THETA
PHI - 0.

THETA IN DEGREES

NORMALIZED ANTENNA GAIN VS. THETA
PHI - 90.

THETA IN DEGREES
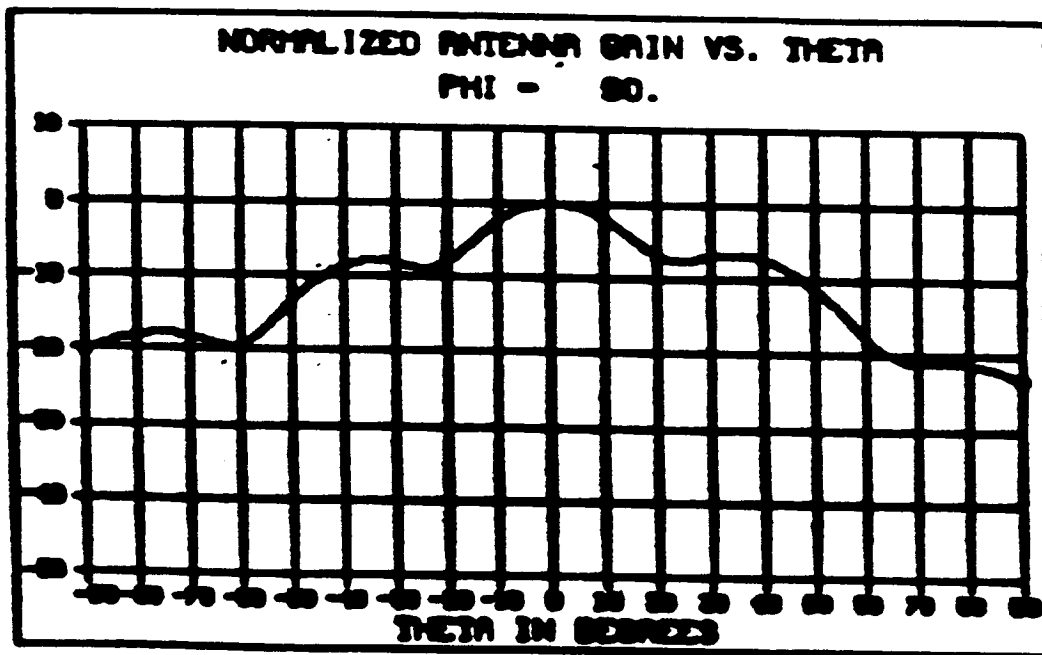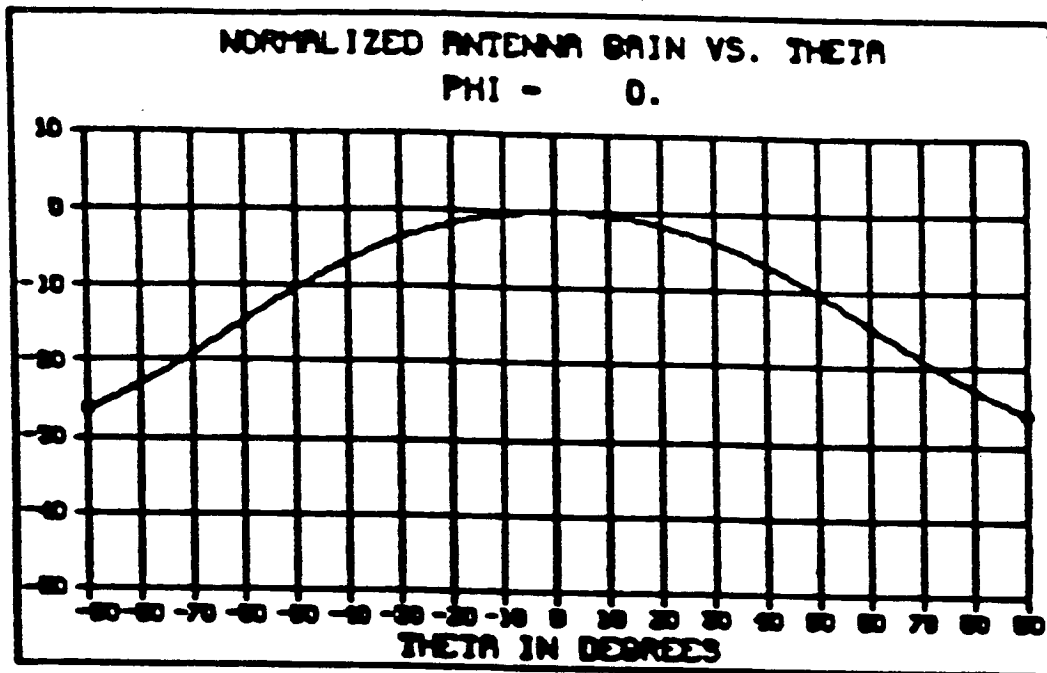
3 HARD FAILURES, SOFT FAILURES

2 X 4 ELEMENT   S-BAND

FIGURE 6

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

University of Houston

INITIAL PLANETARY BASE
CONSTRUCTION TECHNIQUES AND MACHINE IMPLEMENTATION

Prepared by: William W. Crockford

Academic Rank: Research Associate

University and Department: Texas A&M University

Texas Transportation Institute

NASA/JSC

Directorate: Engineering

Division: Advanced Programs

Branch: Systems Definition

JSC Colleague: Ann L. Bufkin

Date: 8 August 1986

Contract : NGT-44-005-803 (University of Houston)

INITIAL PLANETARY BASE
CONSTRUCTION TECHNIQUES AND MACHINE IMPLEMENTATION

William W. Crockford
Research Associate
Texas Transportation Institute
Texas A&M University
College Station, Texas 77843

Conceptual designs of (a) initial planetary base structures, and (b) an unmanned machine to perform the construction of these structures using materials local to the planet are presented. Rock melting is suggested as a possible technique to be used by the machine in fabricating roads, platforms, and interlocking bricks.

Identification of problem areas in machine design and materials processing is accomplished. The feasibility of the designs is contingent upon favorable results of an analysis of the engineering behavior of the product materials. The analysis requires knowledge of several parameters for solution of the constitutive equations of the theory of elasticity. An initial collection of these parameters is presented which helps to define research needed to perform a realistic feasibility study.

A qualitative approach to estimating power and mass lift requirements for the proposed machine is used which employs specifications of currently available equipment from various manufacturers.

An initial, unmanned mission scenario is discussed with emphasis on (a) identifying uncompleted tasks which necessitate manned follow-up missions, and (b) suggesting design considerations for vehicles and primitive structures which will use the products of the machine processing.

The period of research was 16 June - 8 August 1986. The use of names of manufacturers does not constitute official endorsement of such products or manufacturers by NASA or any U.S. government agency.

- - - - - - - - - - - - - - -

NASA Colleague: Ann L. Bufkin ED2 X2536

INITIAL PLANETARY BASE
CONSTRUCTION TECHNIQUES AND MACHINE IMPLEMENTATION

INTRODUCTION AND PROBLEM DEFINITION

Planetary base construction will involve multiple missions due to mass and volume lift requirements from earth. This paper describes a concept for an early unmanned mission which will accomplish the initial tasks of base construction. The mission concept includes, as a key element, the conceptual design of a planetary materials processing machine which will accomplish the initial construction tasks. The machine is intended to produce bricks, roads, and platforms using materials local to the planet. The machine subsystems should be modular in the sense that new technologies which appear before launch can be implemented with a minimum of additional design effort.

Monetary costs of the machine are not directly addressed in this paper. However, the costs in terms of mass lift, operating power, and planetary resource utilization are discussed briefly.

The words "soil" and "regolith" are used interchangeably in this report. The primary emphasis is on lunar base applications because more data is available, it will act as a stepping stone to the other planets, and it is potentially more harsh an environment in which to test ideas and technology than, for example, Mars.

BACKGROUND CONSIDERATIONS

Chemical Processing. The conservation of planetary natural resources is an issue which must be addressed very early in the planetary infrastructure development program. For example, because water is so vital to human presence on the planets, it seems imprudent to make structures of concrete even on Mars (and even using water reducing agents as suggested by Young in [46]) where water, atmosphere, and non-zero relative humidity exist. The carbonation curing mentioned by Young [46] may hold some promise, but is not discussed in this paper. Although it is true that

recent advances in cements, water soluble polymers, and metal and polymer fibers have resulted in excellent concrete products (compressive strengths 30-40ksi (200-300MPa), see Young in [46]), the use of earth-based portland cement concrete technology should be postponed at least until after the establishment of planetary factories which can produce the necessary components of the mix.

Although inorganic polymer chemistry seems to be a promising approach to the problem of concrete type material processing, the progress in this field is apparently confined to linear chains (see Lee in [10]). Lee also briefly reviews a promising technique for the production of high toughness metal glasses which are not discussed herein. In fact, glass may be considered to be an inorganic polymer [67] and is considered as an option, in this paper, for structural material.

Urethane foamed plastic stabilization of lunar soil simulants has resulted in unconfined compressive strengths on the order of 4-5ksi (27.58-34.48MPa) [48]. However, the technique was primarily studied using soil grouting techniques. Extension of the testing to vacuum environments led to problems with the stabilization procedure [49]. Phenolic resins were also tested with unsatisfactory results [49]. Problems with conventional stabilization techniques using stabilizers such as portland cement, foamed plastics, resins, and asphalt products should not come as a surprise if one considers vapor pressure in the analysis. These techniques may be successful on Mars but should not be expected to perform flawlessly on the moon.

Based on qualitative considerations and experimental results, chemical processing of materials for structural purposes was eliminated from consideration.

<u>Passive and Semi-active Mechanical Processing.</u>
Two concepts are of interest here: (1) semi-active techniques such as controlled rock fracturing for shaping building stones or soil moving and placing, and (2) passive techniques such as simple building rock recovery and replacement or adaptation of existing planetary crust formations (e.g. craters, lava tubes). Of these techniques, only soil moving and placing and

adaptation of existing formations are techniques for
which the machine is designed. Controlled rock
fracturing and building stone recovery and placement
are tasks which are too time intensive and which
require too much articulation on the part of the
machine.

Unperfected Methods of Processing. The use of
lasers and microwaves for rock fragmentation by
differential heating of minerals within the rock has
been under study by the Bureau of Mines. The use of
explosives in a vacuum for rock fracturing has also
been studied (see Podnieks and Roepke in [46]). These
techniques are not discussed in this paper.

For the production of construction materials such
as "bricks", microwave processing is a very promising
technology (see Meek et al. in [46]). The total
energy requirements are much lower than those of
conventional heating techniques. Meek et al. have
used 2.45GHz ultra high frequency (UHF) microwaves to
induce diffusion bonded ceramic-glass-ceramic
junctions. Waves of this particular frequency couple
well with ilmenite inducing the necessary initial
temperature rise. While this technology may very well
become the solution to the power requirement problem on
the processing machine, several unanswered questions
have, unfortunately, precluded much more than a cursory
discussion of the technology in this paper. Very
little information on this technique has been published
[PC-13] with the most current and informative article
being that authored by Meek et al. in [46]. Some
questions of interest follow.

(1) Do microwave processed materials have better
engineering properties (e.g. strength, toughness) than
those processed with conventional methods? A
qualitative assessment based on inferences by Meek et
al. [46] would indicate an answer to this question in
the affirmative. However, quantitative information is
needed to confirm this supposition.

(2) Are coupling agents at this frequency too
valuable or scarce to be relied upon for extended
usage? The abundance of lunar ilmenite is generally
less than 2% and may be a valuable source of Fe, Ti,
and O [69]. Mars materials contain valuable water [53]
which will couple.

(3) Is the variability of coupling agent presence over the surface too great to allow product uniformity? Is it a simple matter to identify variations in coupling agents and to adjust the wavelength to couple with a different agent?

(4) Meek et al. [46] state that the ilmenite in an ilmenite-rich basalt couples first causing a temperature rise which, in addition, is sufficient to cause the basalt to couple. Can this "domino coupling" effect be expected to create strong, diffusion bonding in any regolith (i.e. not only in ilmenite-rich basalt)?

In Situ Melt Processing. The Los Alamos Scientific Laboratory (LASL) perfected a drilling technique in 1976 which utilizes simple ohmic heating in a penetrator which creates a dense glass lining around the hole as it drills (see [4], [35], or Rowley in [46]). Because of the relatively complete nature of the research, development, and documentation of this process, this technique was chosen as the process of primary interest. The glass product of this process has higher density, and higher compressive strength than the parent materials (see Rowley in [46]). The process is apparently equally effective regardless of soil or rock composition. There are disadvantages with both the process and the product which are discussed later in this paper.

MATERIALS CONSIDERATIONS

Lunar Materials. Table 1 contains some of the parameters for lunar materials and glass products which are necessary for solution of the constitutive equations of the theory of elasticity and for the solution of other equations used for terrain-vehicle system calculations. Mitchell et al. [51] recorded moduli of subgrade reaction which indicate that insensitive structures may be successfully placed on foundations made of the in situ lunar material. However, the sensitivity of the modulus of subgrade reaction is in question even in the best of circumstances (Horonjeff in [68]). For sensitive structures (e.g. observatories [38]), Mitchell et al. [51] suggest that burying footings deeper where the lunar soil is more dense (which could be done with a

rock-melting penetrator) or compacting the construction
site may be desired to reduce settlements to tolerable
levels. Mitchell et al. also found the soil at
approximately 4-8in (10-20cm) depth to be, in general,
at a very high density. The density distribution with
depth is approximated by [51]:

$$\rho = \rho_o + k \ln(z+1) \qquad (4.1)$$

where ρ_o is approximately 1.27g/cc, b=0.121, z is in
cm, and ρ is in g/cc. The density varies considerably
at the surface on the scale of approximately 3-6ft
(1-2m) laterally.

Table 1. Approximate Lunar Properties.

Parameter	Value(s)	Source
g_o	0.167g	[15]
k	N/A	
α	2.5E-5/degC	[69]
ρ	0.87-1.93g/cc	[51]
Tm	1400degC	[69]
SG	2.9-3.24	[16]
e	0.67-2.37	[51]
c	0.1-1.0kN/m**2	[50, 51, 16, 17]
ϕ	28-50deg	[50, 51, 16, 17]
ksg	800-1600kN/m**2/m	[51]
β	N/A	

As will be illustrated later in this report, the
machine concept addresses the problems of compaction,
removal of soil to higher density depths, and a method
of making the density of the surface layer more
homogeneous from point to point. Modification of the
gradation curve is not a primary purpose of the
machine. The lunar grading curve and soil

classification indicate a well-graded silty sand to
sandy silt (SW-SM to ML in the Unified system [51]) and
further modification to the gradation is not deemed
necessary or desireable by this author. However, the
machine does perform a crushing function as part of the
preprocessing of soil intended for brick production.
This crushing is simply a method of insuring a maximum
desired particle size for the brick, given any soil
input.

The lack of a lunar atmosphere, and, in
particular, the lack of water vapor pressure results in
much lower crack speeds (at the same stress intensity
factors) than those reached at higher water vapor
pressures. Alternatively, one could consider the
stress intensity factor required to attain a given
crack velocity to be significantly greater in the lower
pressure environment [61] as shown in Table 2.

Table 2. Fracture of Lunar Analogue Glasses.

Crack Velocity	Water Vapor	KI
1E-5m/s	10 Torr	21.2N/(mm**1.5)
1E-5m/s	0.001 Torr	25.3N/(mm**1.5)

Lunar Products and Terrestrial Analogues. In situ
melt processing of lunar materials will produce a glass
which may have questionable strength properties.
Specifically, cracking may be a problem. The cracking
may be a manifestation of residual stress problems or
thermal stress induced fatigue. However, cracking may
not be as serious as it first appears, especially in
the absence of corrosive agents such as water vapor.
If angular "aggregates" result from cracking of the
glass, roads and platforms may still perform acceptably
due to aggregate "interlock". In the case of bricks
and melt-tracks, however, performance may be seriously
impaired by cracking.

ORIGINAL PAGE IS
OF POOR QUALITY

Table 3. Approximate engineering properties.

Item	Tuff Glass	Dry Bldg Brick	PCC	Silica Glass	10%SC	Al2O3	18Ni Steel
E(GPa)	7	N/A	14	70	2.5	350	207
ν	0.3	N/A	0.18	N/A	0.15		0.29
k(W/mm/degC)	7E-4	6E-4	10E-4	12E-4	N/A	290E-4	150E-4
α (/degC)	0.69E-5	0.9E-5	1.3E-5	0.05E-5	N/A	0.9E-5	1.6E-5
ρ (g/cc)	2.23	2.3	2.4	2.2	1.8	3.8	7.93
σ_c (MPa)	50***	55	34	137	N/A	2000	N/A
σ_y (MPa)	1	N/A	3	10	1.1**	172	1930
K_{Ic} (MPa m**0.5)	0.77*	N/A	0.87	0.75	0.15	4.2	94
J_{Ic} (N/mm)	N/A	N/A	0.035	N/A	0.0085	N/A	N/A

*	Estimated from [61] lunar glass analogue
**	Indirect tension [24]
***	Hollow cylinder test [55]
N/A	Not Available in sources referenced

The glass lining of the rock-melt drilling process has been characterized as transversely isotropic (cylindrical coordinates) by Nielsen et al. [55]. The axial and tangential material properties were found to be equal (E=8GPa, 20GPa, ν=0.34, 0.24 at 0 and 50MPa confining pressure, respectively). The radial properties were found to be slightly different (E=6GPa, 14GPa, ν=0.26, 0.16, at 0 and 50MPa confining pressure). In Table 3 and Figure 1, a comparison of important engineering parameters extracted from various sources [12, 24, 31, 36, 43, 47, 55, Rowley in 46, 61, 70] is presented.
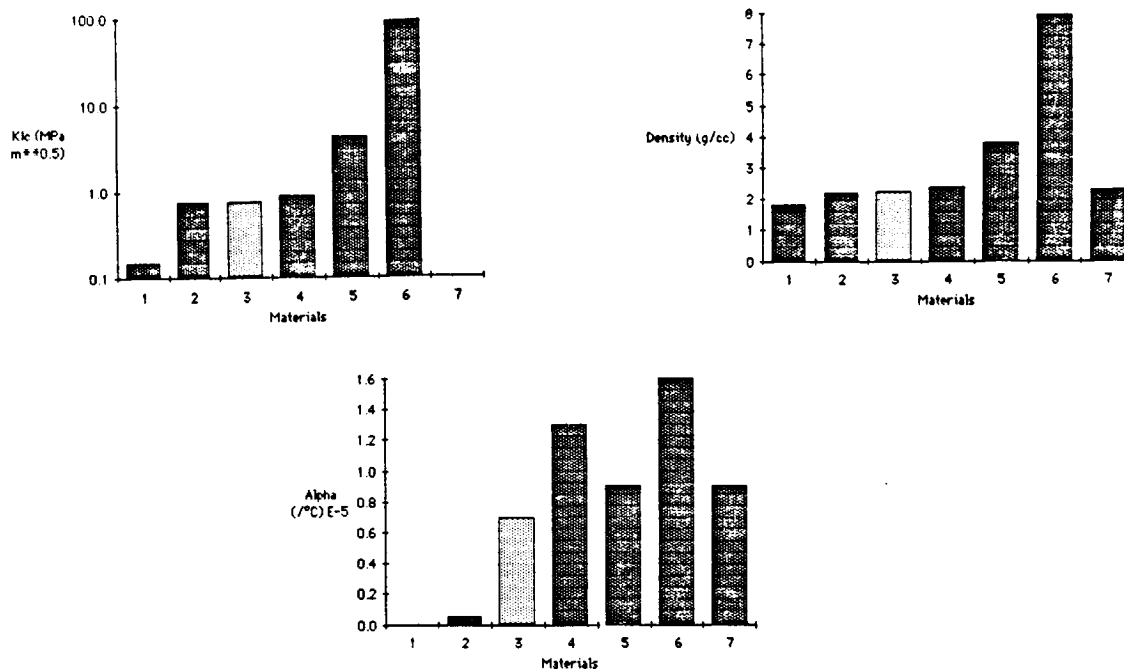
1=SOIL CEMENT, 2=SILICA GLASS, 3=TUFF GLASS, 4=PCC
5=AL2O3, 6=18NI STEEL, 7=DRY BUILDING BRICK

Figure 1. Graphic presentation of selected parameters
  from Table 3.

     It is difficult to compare many of the values in
Table 3 and Figure 1 because of the different test
methods involved. However, it is useful to note that
lunar glasses, fused silica, and a rather ordinary
plain portland cement concrete (PCC) have fracture
toughnesses of the same order of magnitude. An order
of magnitude study also indicates that glass from rock
melting has a compressive strength comparable to both
plain PCC and ordinary dry building brick.

     Martian Materials. Little information is
available concerning engineering properties of the
martian regolith. However, some properties have been

approximated by studying footpad penetrations, descent
engine induced surface behavior, and surface sampler
data [53, 54]. In Table 4, some of the available
parameters of interest are presented.

Table 4. Approximate Martian Properties.

Parameter	Value(s)	Source
g_m	0.38g	[53]
k	N/A	
α	N/A	
ρ	0.6-1.69g/cc	[53]
Tm	N/A	
SG	N/A	
e	N/A	
c	0.01-14kN/m**2	[53, 54]
ϕ	18-45deg	[53, 54]
ksg	207-1600kN/m**2/m	estimated [51, 53]
β	39deg	[53]

The problem of working with martian soils may, at
first, seem much simpler than working with lunar soils
because of the presence of water as a processing aid.
However, several factors make the martian base at least
as challenging as the lunar base.

(1) Minimization of the nonrecoverable use of
water is mandatory.

(2) The presence of water and high relative
humidities [25] delete the advantages for glass
utilization present on the moon. That is, nonzero
water vapor pressure tends to act as a mechanism for
enhancement of stress corrosion cracking. Carbon
dioxide atmospheric effects are not studied in this
paper.

(3) The occurrence of freeze-thaw cycles [53] is virtually assured.

(4) The presence of montmorillonitic clays [53] may be detrimental to structural materials during wet-dry and freeze-thaw cycles. On Earth, these type clays often exhibit dimensional instabilities in the presence of wet-dry cycling.

Air entraining agents are often used to help alleviate the problems associated with freezing and thawing. The agents insure a discontinuous pore system made up of very small bubbles by acting as surface-active agents. The process used for making structural materials out of martian regolith may require (a) removing all water from the system, except perhaps for tightly bound water interior to the double diffuse layer, and/or (b) waterproofing the component. It may be possible to accomplish freeze-thaw protection using a combination of sintering and rock melting techniques, but the protection may come at some unknown cost in terms of an increased susceptibility to stress corrosion cracking.

INTERIM SOLUTION

The materials processing requirements are temporarily met by a conceptual design which allows removal of soil from the surface down to a depth which gives a relative density [51] of 90% (i.e. approximately 20cm). Manufacturing of bricks may be done as the soil is piled in windrows during the removal operation. The density at 20cm depth is on the order of that of stabilized base materials. Compaction to 95% relative density is then accomplished by rollers. Pressure applied by pad feet is arbitrarily set at 23kg/cm**2 (e.g. Caterpillar model CP323), single lift of 10cm which requires approximately 27.6kg/cm**2 to give proper compaction. If the pad foot could be made large enough (i.e. if the vehicle were heavy enough), the operation could be completed in one pass. Realistically, however, multiple passes would be required. Excavation and melting of shallow trenches would be the next operation for the making of a glass track or "rail" system for follow-on vehicles. The vehicles could be maneuvered in the trench-rails by mobile controls (e.g. [PC-10]). The "road" would be

4-12

made straight and level by using a laser system (e.g.
[PC-3]) or some other navigational aid.

Utilization of rock melting or even hot press
sintering techniques will probably require on the order
of 4.5kW power. Using 5.14W/kg [8] for a conservative
radioisotope thermoelectric generator (RTG) power to
weight ratio results in an 875.5kg power unit. In
Figure 2, power versus weight is illustrated for
various equipment (88 construction machines of 4
manufacturers).
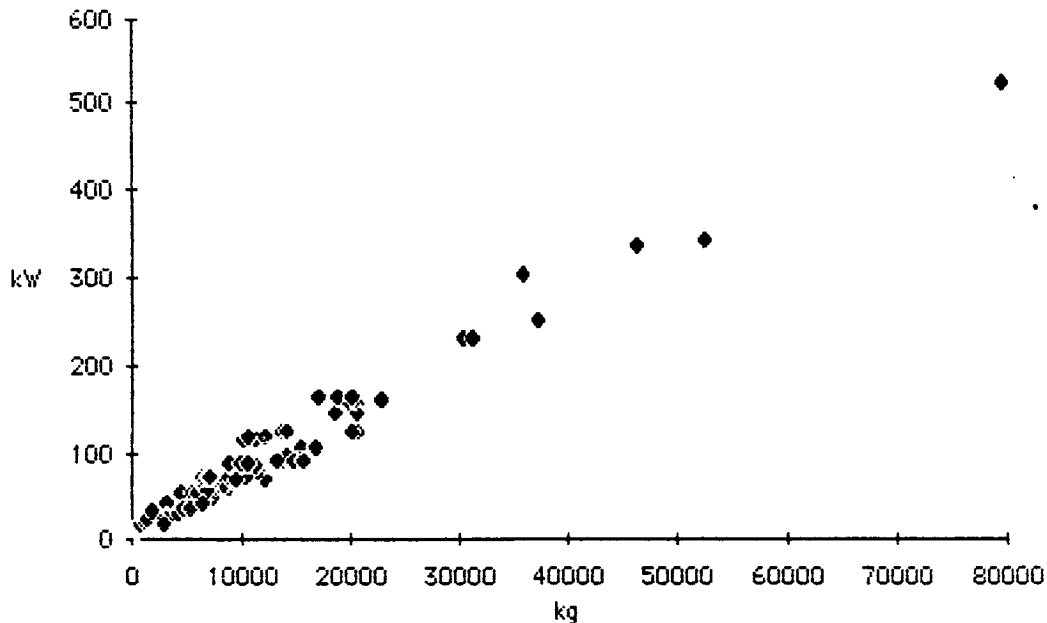
Figure 2. Power versus weight for various construction
    machinery. Y=0.11(X**0.72), R-Square=0.93, N=88.

As an illustration of general power versus weight
requirements, assume the planetary soil has an earth
unit weight of approximately 1600kg/m**3 in the loose
state and we desire to push 0.75m**3 of the soil (e.g.
John Deere 650). Using a conservative coefficient of
traction factor of 0.5 for loose earth, dry sand and
clay loam [18], the minimum total vehicle weight should
be on the order of

1600*0.75=1200kg
or 456kg on Mars

456/0.5=912kg total machine weight on Mars

Using the regression model of Figure 2, a power
requirement of 14.8kW is suggested. Therefore, the
machine would be underpowered by approximately 10kW for
soil working purposes in terms of existing
manufacturer's equipment.

Using the Caterpillar models 815 and 825
compactors for the sake of example, it is seen that the
weight of engine and fuel is approximately 65% of the
weight of the vehicle. This would imply that an RTG
weighing 1560kg on earth and producing on the order of
25W/kg will allow successful design of the
aforementioned underpowered vehicle. The SP-100
program (see French in [46]) gives a glimmer of hope
that this power capability is attainable. This
analysis leads us to an estimate of the size
requirement for the machine on the order of a John
Deere model 675 (approximately 9 cubic meter volume,
3.2m long, 1.6m wide, and 1.8m high). This vehicle
size is also on the order of magnitude of the Boeing
LRV [22].

Comparisons using the regression relationship of
Figure 2 may be somewhat qualitative when one considers
that (a) the regression is for gasoline and diesel
engines manufactured in discrete power ranges, not for
electric power supplies and motors, (b) it is suspected
that design procedures for common earthworking
equipment has not really attempted to maximize
production while minimizing both weight and power
requirements, (c) there is a very small difference
between the weight of the power unit and the weight of
the complete machine in the case of the 5.14W/kg RTG
example, and (d) the RTG and SP100 units contain their
own fuel system while the fuel weight for the engines
of earth construction equipment is impounded in the
vehicle operating weight.

Bricks will be made from the top 20cm of soil
which was removed in the original clearing operation.
The windrows of soil would be removed from the surface
and transported by a belt system to a crusher and sieve
before entering the mold for hot pressing, sintering,
and/or melting. Gravity flow of the material through
bins is not necessarily straightforward, but has been
studied [56]. The question of the desireability of a

hot pressed, sintered brick versus melting to a glass coated or solid glass brick will not be resolved without experimental study. However, if the glass layer is annealed properly, and good bonding with and compaction of the interior of the brick is attained, the glass brick may have two advantages over the standard brick:

(1) Higher density surfaces may allow decreased thicknesses of protective radiation shielding.

(2) The relatively high stiffness glass layer may (depending on the silica content and the effectiveness of the annealing process in combating cracking) result in higher strengths than the sintered material not only because of the inherent material properties but also for the same reason that a testing machine using a steel platen will yield higher concrete strengths than a "brush" platen [47].

Glass coated "bricks" are not an idea new to the natural lunar environment. Several glass coated rock specimens have been studied. Some of the samples have excellent strengths and impact resistance while others are very fragile [PC-2]. The difference in impact resistance of the glass layers of different samples is thought to be caused by thermal treatment over geologic time [PC-2]. An effort should be made to understand the cause of the difference in glass layer impact resistance and hypothesize how the process can be accelerated for use as a construction material processing technique. Elastic wave velocities are available for some samples such as lunar sample 60015 [58], but other strength and material properties are needed for this and similar samples.

In Situ Rock Melting. Both lunar and martian soils contain volatile elements [15, 69] which can be driven off by heating. It has been noted [35] that the evolution of gas during melting may cause voids in the glass. This problem was solved in the LASL research program by increasing the melt pressure at the glass-forming section. Apparently, attainment of high pressure ahead of the penetrator stagnation point is mandatory (see LASL Mini-review 75-2 in [57]). Outgassing of lunar soil was also observed in compression and shear testing by Carrier et al. [17]. The penetration rate is limited by the heat flux that

can be provided at the leading edge or stagnation point [35].

In the brick manufacturing process proposed for the planetary materials processing machine discussed herein, maintenance of pressure in the mold using the limiting case of the flat plate melter will not be a significant problem. However, in the case of the track-melt process discussed later in this paper, confining pressure may become slightly problematic. It is suspected that the problem is not catastrophic.

The thickness of the melt layer in a given soil or rock can be controlled by causing electric heating current to flow directly through the rock melt layer, by using a long conduction heating section, or by introducing pellets of material to be melted [2]. Since the glass layer is typically of the order of 4-15mm thick in the LASL studies [2], control of the glass thickness may be required. Specifically, thicker layers may often be required. A new concept specifically formulated for the planetary materials processing machine is introduced to identify a procedure for thickening a glass layer. It is speculated that annealing of a horizontal glass plate (or hole lining) may be accomplished simultaneously with increasing the thickness of the plate as follows:

(1) Form the initial glass lining.

(2) Add loose raw material on top of the plate. The amount of added material may govern the depth of annealing.

(3) Melt the new material so as to anneal the lower portion of the initial glass lining.

In this manner, heat treatment of the glass lining may proceed from the outer lining material toward the melter body. The last layer would have to be annealed in a separate operation.

The glass layer produced by this technology will, undoubtedly, have cracks similar to the radial cracks observed by Nielsen et al. [55]. It is not known at this time whether annealing, high pre-melt in situ porosity, or mineral composition of the materials will allow production of an uncracked lining. One should

expect to use fracture mechanics as the method of failure analysis because preexisting sharp flaws are virtually inevitable. Theoretically, a very thin glass layer is less likely to break due to thermal stresses [64] and may decrease the tensile stress (even resulting in compression) in the top layer of a three layered semi-infinite half-space [71]. The layered elastic result is, of course, dependent on modular and thickness ratios of the two top layers (top/middle) [71]. Given the existence of cracks in the glass layer and that the tensile stresses in a stiff top layer may also be reduced by increasing the thickness of the top layer in relation to the middle layer, maximization of the top layer thickness is considered prudent. The desire for increasing thickness should decrease as the modular ratio decreases. Shear and deflection are also affected by changes in these parameters and must be considered in the analysis.

Theoretical analyses using features of the AYER finite element code [see 35] and experimental verification are needed to transform the speculative concept of simultaneous annealing and thickening into a viable process.

The penetrators used in the LASL study usually employed ohmic heating [4]. At least on the moon, electron-beam heating [4] may be an interesting alternative because of the availability of a vacuum environment.

Power requirements of the LASL penetrators in the 2-6kW range are plotted in [34] as the abscissa in plots describing the calculated performance of a double-cone consolidating penetrator (approximately 75mm diameter, 180mm long). In general, penetration rate (approximately 0.05-0.17mm/s) and surface temperature (approximately 1700-2000K) both increase with total power, while the minimum required thrusting force decreased at a decreasing rate (approximately 0.6-0.1kN). The power requirements of the LASL penetrators are capable of being satisfied by RTGs [8] or SP100s (see French in [46]).

Planetary Materials Processing Machine. The machine concept was formulated by defining its mission and searching for existing technology which could be adapted to the mission. Three primary machine tasks

were identified: (1) clearing and grading (i.e. bulldozing), (2) compacting, and (3) brick and wearing surface fabrication. The capabilities to fracture, rip, mill, adjust compaction drum weights, provide vibratory compaction, and/or vertical articulation of the drums (e.g. Ingersoll-Rand model LF-450) were considered, but eliminated from the design for two reasons: (a) cost in terms of machine complexity and expected reliability, and (b) clever choice of the initial landing sites can be used to eliminate the necessity for these additional capabilities. Steel track beadless tires [18] and track systems, including new low ground pressure tracks [PC-7] were considered, but eliminated from the design because of the desire for compaction. Tracks could be placed over the drums after the compaction operation but this operation would probably require a manned presence. In addition to the required processing capabilities, the machine is intended for an unmanned mission and must have good mechanical reliability (which is much more important than production rate for this mission), excellent navigational capabilities, and remote control/data analysis functions. The resulting machine concept is basically a static compaction machine with a bulldozer blade and a brick making unit. The machine is vaguely reminiscent of equipment such as the Ingersoll-Rand models SPF-56B, DA-28, or the Caterpillar models CP-323, 815B, 816B, CB-214, combined with the Boeing LRV. A diagram of the chassis and drum wheel system is presented in Figure 3.

The size of the machine has been scaled down in this paper in order to lower the power requirement. The overall size of the machine and especially the drum size should be increased to the maximum size possible at launch time. A lander vehicle will house a data/remote control link with Earth and a Laserplane [PC-3] type transmitter. The laser receiver is placed on the rear left corner of the machine chassis. The melter of Figure 6 is mounted at the bottom of the chassis between the wheels, is capable of vertical articulation, and is positioned with the longitudinal axis parallel to the y-axis of Figure 3 with the "bullet" nose pointing in the negative-y direction. Redundant systems should be available for all primary systems except the power generator (e.g. the machine should be able to carry out its mission even if part of the four wheel drive capability is lost).
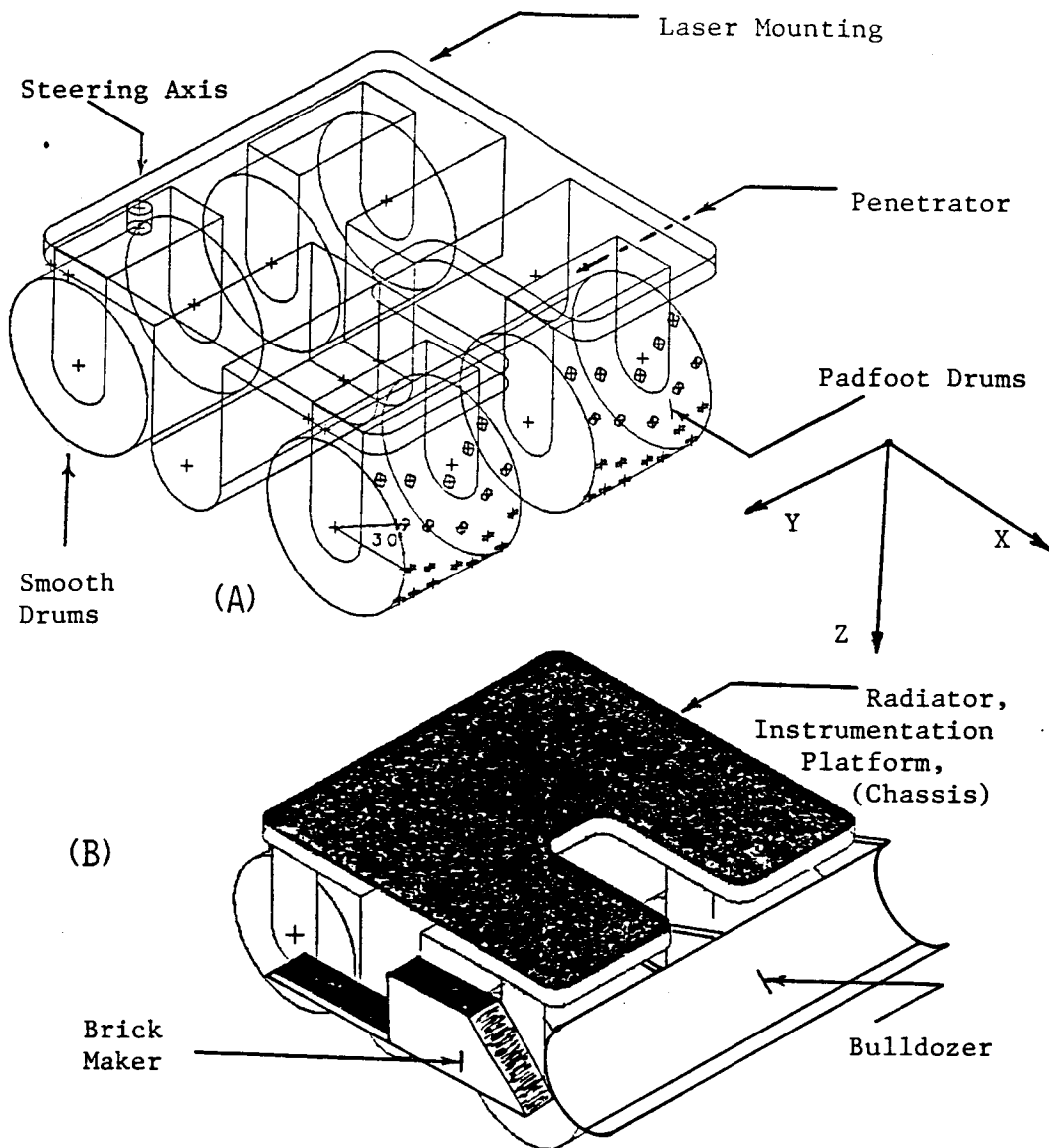
Figure 3. General concept of the machine:
   (a) transparent, (b) solid model with attachments.

The concept of the bulldozer blade and the brick making
system is illustrated in Figure 4.

(A)

Crusher

Sieve

Special
Purpose
Modules

Mold

Soil Pick-up

(B)

Grooved Belt
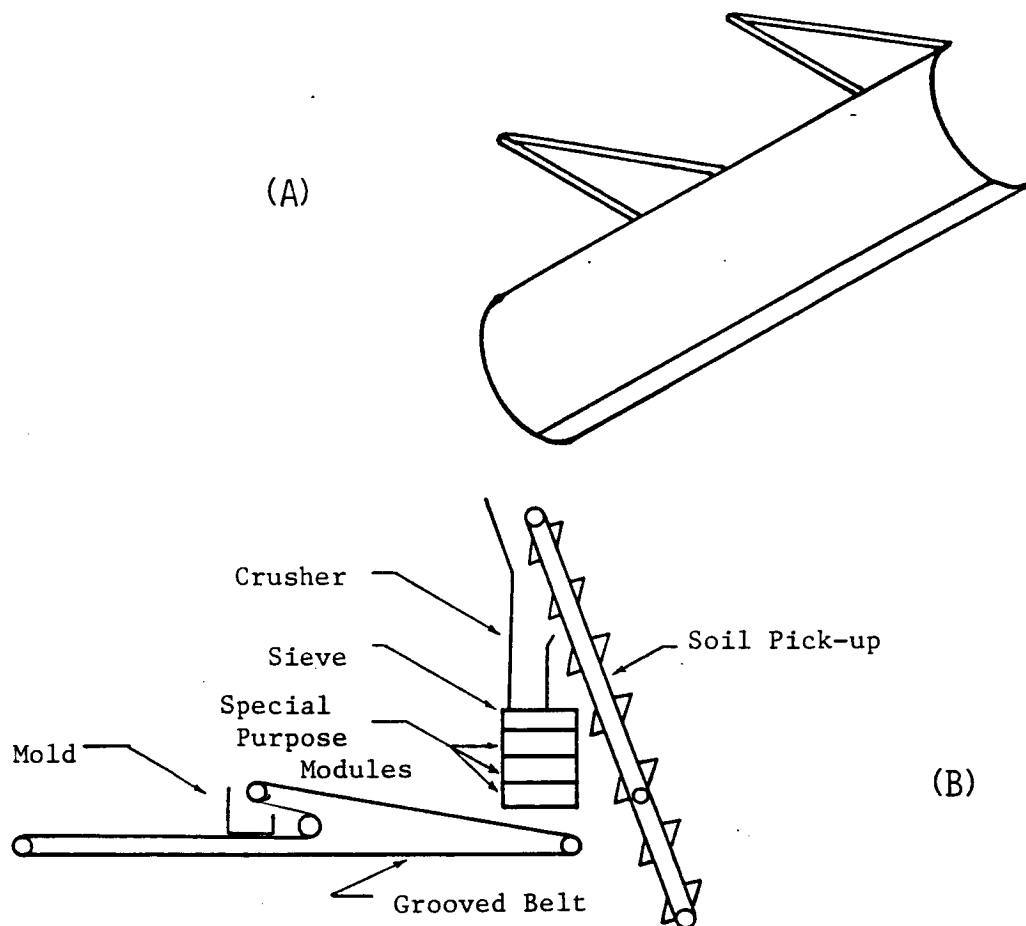
Figure 4. Concepts for (a) blade, and (b) brick maker.

In general, systems which are normally driven by
hydraulic actuators on earth machinery will be driven
by electric servomotors and jackscrews. The machine
subsystems are categorized as follows:

Chassis:
        Thermoelectric power generation and transmission
        Heat disposal
        Navigation (primarily LASER)
        Track melter

Drum wheels:
        Steering

Propulsion

Blade:
        Servomechanisms

Materials processing unit:
        Belt system
        Crusher
        Sieve
        Raw material preheater
        Brick producing mold
        Brick cooling system

The drum wheels rotate 90 degrees either direction about an axis parallel to the z-axis in Figure 3. This rotation is necessary for three operations: (1) generation of windrows is accomplished by angling the wheels (which also allows immediate pickup of raw soil for brick production if properly designed), not by angling the dozer blade as is typically done on earth based machines, (2) final smoothing of a surface layer can be done by turning the drums to the 90 degree position so that only the smooth drums contact the surface while the padfoot drums are off the smooth surface, and (3) generation of the melt tracks discussed later in the report require the full 90 degree rotation. Discarded designs which also utilized the 90 degree rotation capability included the use of a flat, heated dozer blade for the construction of vertical walls and a microwave concept for binding individual bricks to each other (somewhat like slipforming the mortar in a brick wall). Since the machine basically has no suspension system (i.e. no vertical articulation of the drums), machine stability considerations dictate that the dozer blade is required not only for making smooth and level platforms, but also for leveling ahead of the machine. The center portion of the chassis is designed to house the major, heavy components of the power unit allowing the center of gravity to be lowered.

The machine is envisaged as being able to make only one brick at a time. An "interlocking" brick design for use without mortar is illustrated in Figure 5. A heated mold for the brick is used. The "sides" of the mold are stationary, while the top plate applies

pressure (directly opposed by the heated bottom plate) during heating. The bottom plate is hinged so that the top plate can be used to extrude the finished brick. The process of making the brick is estimated to take approximately 1-3 hours (estimated from [3]).
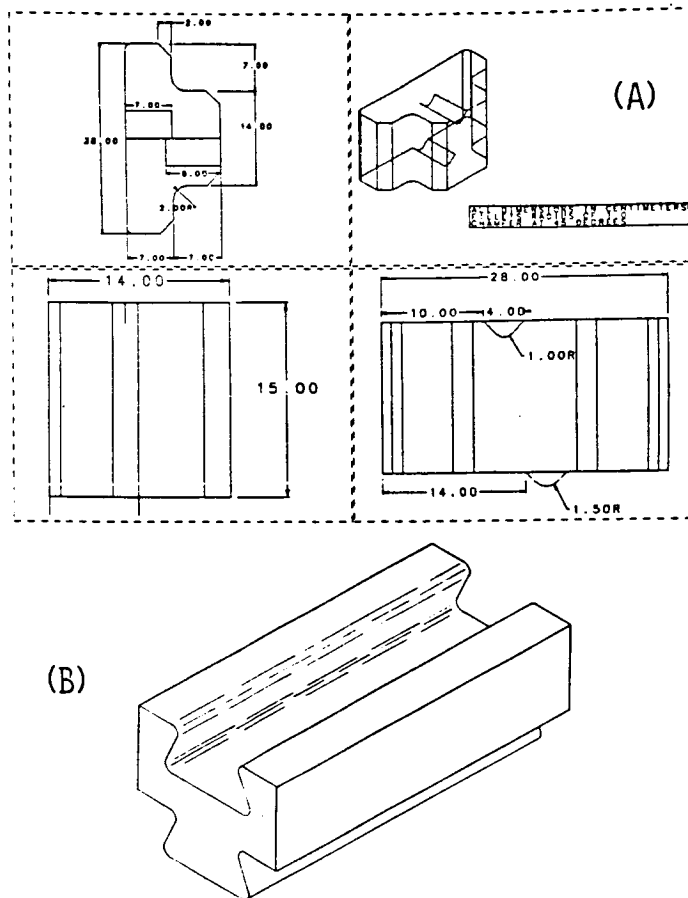
Figure 5. Mortarless brick design: (a) T shaped, (b) Dove tail.

No detailed overall machine design has been accomplished. Problems which must be addressed before detailed design can be done include, but are not limited to:

(1) Brickmaker: Is the problem of gas evolution serious, or very simply handled by the pressure induced by the top plate of the mold? How critical is the fit between the top plate and the sides of the mold? How will cooling be controlled? Is complete melting to glass desired, or is a standard hot press technique actually better?

(2) Drums: Padfoot design (e.g. height, cross sectional area, and geometric design) must be optimized for the expected soil conditions and machine weight. Typical static pad foot pressures of $20-50kg/(cm**2)$ are attained with earth machinery. Lighter machines and smaller cross sections of the pads will result in the requirement for more passes to attain desired compaction. The smooth steel drums were originally intended to be heated for rock melting of large platforms, but the idea was discarded due to power requirements and due to anticipated problems with the thermodynamics of the melt layer. The smooth drums are retained in the design in order to provide some static compaction in sandy soil and to act as finishing rollers. If desired, the smooth drums may be wrapped with a treaded belt to gain some of the advantages of rubber-tired rollers. Drum diameter will be determined by the soil conditions, vehicle weight, and desired dozing capability.

(3) Blade: The blade has not been designed in this report. Some considerations on the design can be found in [5, 7].

(4) Penetrator: A design of the glass forming section of the penetrator is needed which will address the question of confining pressure. A proposed design is illustrated schematically in Figure 6.

In the pursuit of the solution to the detailed design problem, modules to be added to existing computer programs are proposed. Specifically, bulldozer and melt processing modules should be incorporated into the MSFC LRV analysis program [22]. The modified LRV program should then be incorporated in

the IDEAS**2 package [PC-1]. Extension of the AYER
[35] finite element program to three dimensional
analyses (i.e. beyond the axisymmetric case) should be
attempted at this stage of research. Solution of
several of the equations in the terrain-vehicle
interaction modules requires knowledge of the Bekker
soil value parameters [7] which are given for lunar
type soils in Table 5.
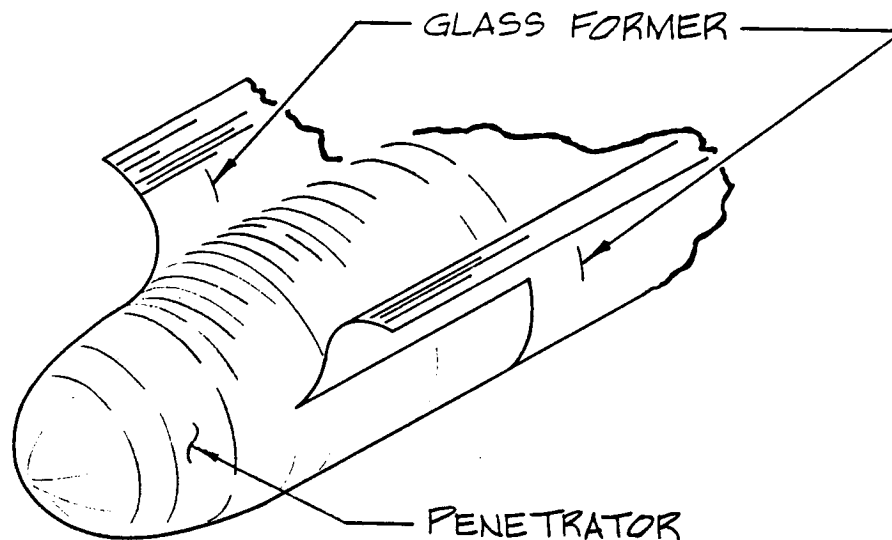
Figure 6. Schematic of track melting penetrator.

Table 5. Bekker soil values [22].

n	k_ϕ (lb/(in**(n+2)))	k_c (lb/(in**(n+1)))
0.92*	6.26	1.13
1.0**	3.0	0-0.4

* sand, basalt
** estimate for lunar soil

The resulting computer output should allow computation
of the ratio of (power available/power required) versus
total machine mass. It is anticipated that the correct
solution is attained in an iterative fashion when the
ratio is greater than or equal to 1.0 and the total
mass of the machine is large enough to accomodate the

mass and volume of the power supply. The suitability
of Bekker's approach to the locomotion problem has been
challenged in one case [48]. Since Bekker's method is
not completely analytical (i.e. several empirical
constants are required), careful review of the
equations involved and of the statics and dynamics of
the problem are required prior to implementing any
computer codes.

On the other hand, three problems not directly
associated with the machine design can be formulated
and solved in the short term (upon acquisition of the
material properties of the brick) using the finite
element method.

(1) Thermally induced stresses or displacements in
the brick due to surface temperature fluctuations on
the order of 290 degrees C may be calculated in
preparation for failure analysis of the brick.

(2) Stresses or displacements resulting from
opposing distributed loads on the top and bottom of the
brick are needed. The impact of meteorites and moon
quakes (on the order of Richter 4 [69]) on assembled
walls should also be assessed.

(3) Stresses or displacements in the soil volume
ahead of a bulldozer blade are needed.

Incorporation of probability distributions into
the finite element code as a statistical approach to
handling the problem of inhomogeneous, possibly
nonisotropic material should be a long term goal of the
computer code research.

STRESS ANALYSIS REQUIREMENTS

Bricks and Walls. The use of bricks for walls may
be necessary for purposes such as radiation protection
even in the cases where reactors are placed in craters
(e.g. French in [46]). Stress analysis of the bricks
will probably take place in a finite element context
because the geometric boundary conditions will make a
purely analytical derivation difficult at best. The
walls must be analyzed for performance under the
following loads: (a) meteorite impact, (b) thermal

cycling, (c) quakes on the order of Richter 4, (d) degradation by the solar wind; and, for Mars, (e) wind loads [60], and (f) stress corrosion induced by the atmosphere.

An early concept for an igloo shaped structure (shell of revolution) with a very large base diameter to height ratio was originally considered for use as a storage and habitat facility. The concept was to pile material in a mound, roll over the mound with the melting drums, and then excavate the loose material from beneath the glass melt shell (similar to concepts discussed by Khalili in [46]). Although the shell of revolution often allows relatively thin structures to be built, this method was considered to be impractical for several reasons not enumerated herein. In addition, excluding the drum melters during the redesign of the machine has precluded this option. Shell structures will therefore have to be launched from Earth or made using a curved brick design which has not been done in this paper.

Melt Tracks. Geotextiles were considered for use as road surfaces but were eliminated from the design. A concept of a melt track rail system which is of interest is illustrated in the initial base shown in Figure 7.

The melt track system may be useful as a test bed for the following concepts:

(1) Mining car transportation of materials may proceed by means of linear induction motors placed in the soil between the melt tracks and with guidance of the car assisted by laser or other mobile control systems [PC-10]. Propulsion of rather large payloads at 17fps is attained routinely, with much higher speeds attainable for optimized linear induction designs [PC-9]. Pressure applied to the tracks by the cars will be an important factor in the feasibility of this operation.

(2) If the speed of vehicles using the tracks can be increased to a reasonable rate, the system can be used for rapid transit of astronauts from frequently visited work areas remote from protective habitats back to the habitats in the event of harmful solar activity.

C - 2

(3) If the glass melt technique can be perfected, it may even be possible to use a modified form of this concept in mass driver design. The system could be designed for very low friction, thus minimizing the time and distance needed for acceleration.
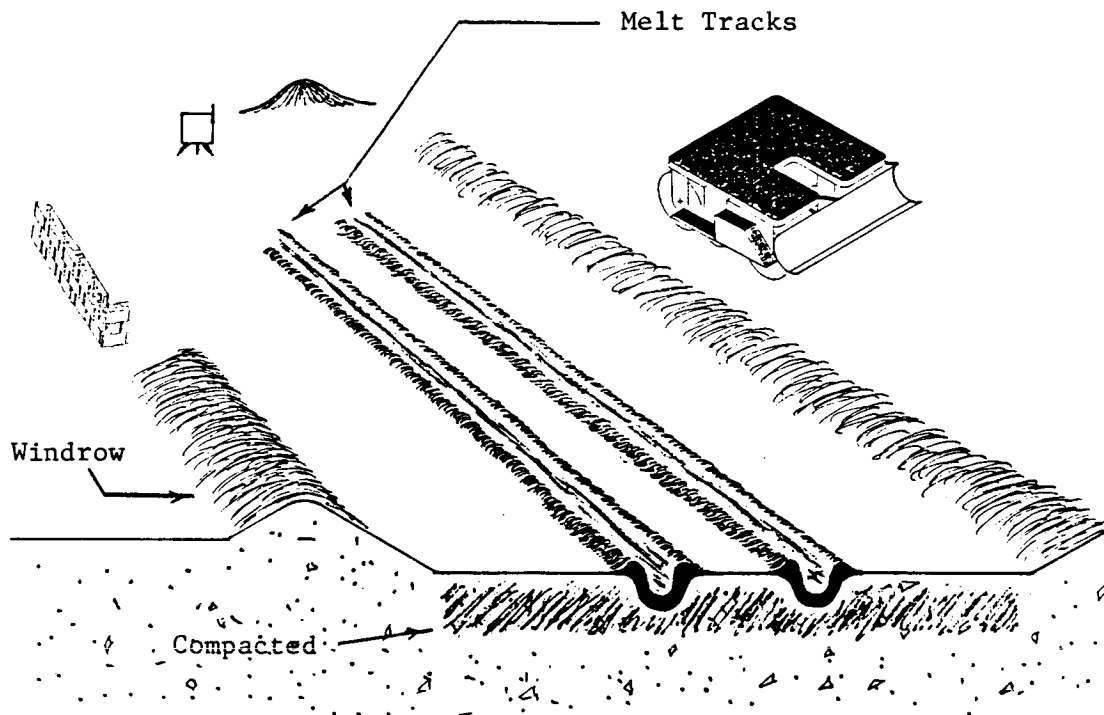
Figure 7. A portion of the initial base.

The three operations mentioned above are nothing more than wishful speculation without a stress analysis of the track system. In Figure 8, the problem is defined.
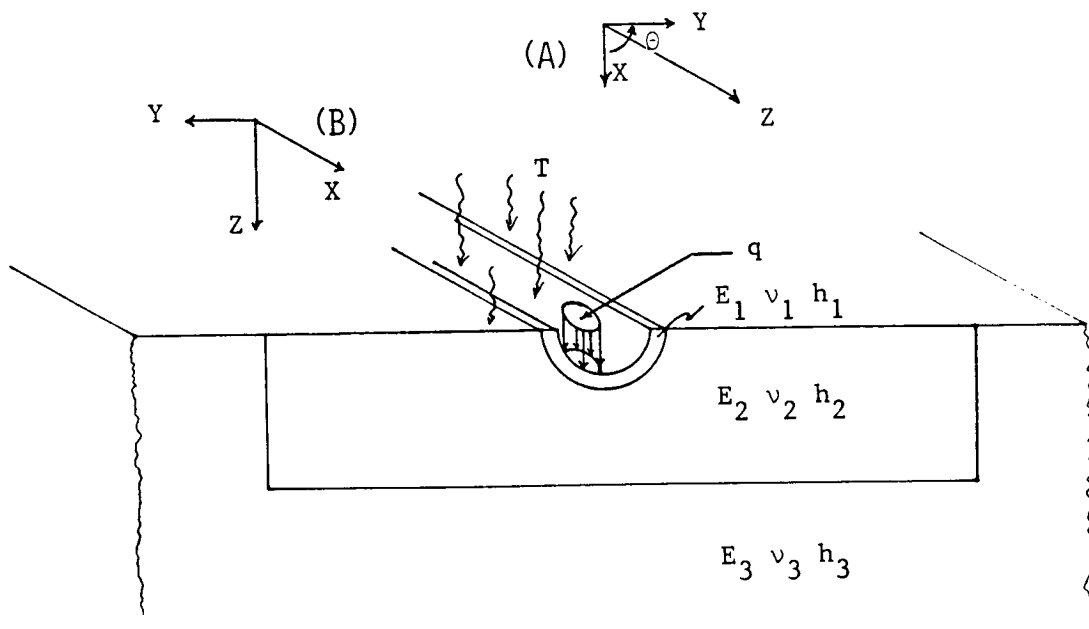
Figure 8. Problem definition: melt track.

An attempt to solve the problem using an analytical
approach to the solution of the stress field in the
glass track may precede a finite element analysis. The
necessity for the liberal use of the principle of
superposition in the derivation must be tempered by a
critical assessment cf whether or not superpostion
applies (e.g. the displacements and displacement
gradients should be small so that the Lagrangian and
Eulerian infinitesimal strain tensors are approximately
equal). As is often the case when some lack of
understanding is present and approximate solutions are
acceptable, linear elastic behavior and superposition
validity are assumed. Returning to Figure 8, it is
noted that the wheel loads can be placed in the
interior of the glass track where St. Venant's
principle tells us that the end effects become less
important, or the load can be placed on an end of the
glass track simulating an expansion, construction, or
terminal joint. The general problem is reduced to
three problems, the solutions of which may be
superposed.

    (1) an axisymmetric thermal effects problem,

(2) a distributed load over a portion of the
boundary of a multiple layered plate with a
semi-infinite bottom layer, and

(3) the penny-shaped crack normal to a boundary,
or the crack in a cylindrical shell.

An example of the general approach to the solution of
the problem follows.

Using axis system (A) of Figure 8, the approach to
problem (1) generally follows the guidelines given in
Timoshenko et al. [64] with the result that
(a=interior radius of the melt layer, b=exterior
radius)

$$\sigma_r = \frac{\alpha E}{1-\nu} \frac{1}{r^2} \left(\frac{r^2 - a^2}{b^2 - a^2} \int_a^b Tr \, dr - \int_a^r Tr \, dr \right) \qquad (4.2)$$

$$\sigma_\theta = \frac{\alpha E}{1-\nu} \frac{1}{r^2} \left(\frac{r^2 + a^2}{b^2 - a^2} \int_a^b Tr \, dr + \int_a^r Tr \, dr - Tr^2 \right) \qquad (4.3)$$

$$\sigma_z = \frac{\alpha E}{1-\nu} \left(\frac{2}{b^2 - a^2} \int_a^b Tr \, dr - T \right) \qquad (4.4)$$

The maximum tangential stress at a free end is modeled
as a beam on an elastic foundation by taking a
longitudinal strip from the shell [64].

$$(\sigma_\theta)_{max} = \frac{\alpha E T}{2(1-\nu)} \left(\frac{\sqrt{1-\nu^2}}{\sqrt{3}} - \nu + 1 \right) \qquad (4.5)$$

The approach to problem (2) begins with the
Boussinesq solution (see [64]) for a load distributed
over part of the boundary of a semi-infinite solid as
modified by Burmister [13, 14] for three layered
systems. Computer implementations of layered elastic
analyses have already been accomplished and can be used
for thickness design of the melt layer, given basic
material properties of the various layers. These
programs sometimes give only the stresses, strains, and
displacements at the layer boundaries. Therefore, an
attempt must be made to correct the stress distribution
[64]

$$\sigma_r = \sigma_\theta = \frac{q}{2} \int_0^a [-2(1 + \nu)z(r^2 + z^2)^{-3/2} + 3z^3(r^2 + z^2)^{-5/2}]r \, dr \qquad (4.6)$$

for multiple layers so that an analytical, continuous form is available within layers to be projected onto the crack plane in problem (3). In equation (4.6), axis system (B) of Figure 8 is used with 'a' being the radius of a circularly distributed load on a flat boundary. It is expected that the load will actually be elliptical and distributed on a curved boundary as depicted in Figure 8.

In the solution to problem (3), it is expected that the Schwarz-Neumann alternating method [see 41] will be used. Basically, the thermal and wheel load stress solutions will be superposed to give the stress distribution at the location of the crack and the stress distribution remote from the crack (i.e. at the surface). This process will establish the initial traction or displacement boundary conditions. The alternating method is then used to generate the cracked body solution for the strain energy density factor.

After all this theoretical work is done, it will be necessary, once again, to apply some sort of statistical method to results. This adjustment is required because of the difference between the real material and an ideal continuum. It is also necessary because no consideration has been given to the interaction of crack tips, nor has consideration been given to possible dynamic effects in this scenario. Alternatively, the influence of multiple cracks may be approached by using a technique based on anisotropy of elastic response resulting from the assumption of or knowledge of a crack density tensor [40].

MATERIAL TEST PROCEDURES

Some of the ASTM test procedures which would be useful in determining the needed parameters for evaluation of alternative products of the machine are identified in Appendix D. Many of the tests mentioned should be used as guidelines for the development of new methods rather than rigid procedures because of the nature of the materials and geometries involved. The most important tests to accomplish in the short term

are those which yield the parameters needed for stress
and displacement field analyses and those which define
failure mechanisms.

It is expected that a maximum principle strain
criterion will be appropriate for many of the product
materials and that the Mohr-Coulomb [see 62] failure
envelope will be quite useful. Failure by fracture of
these materials is expected, but will probably involve
mixed modes and multiple crack interactions. One
approach to the problem of microcracking around a
macrocrack as applied to concrete can be found in [6].

The parameters needed for the penetrator and brick
mold plates are already available through LASL. The
pad feet and drums for the compactor are usually made
of work hardened manganese steel [PC-8] for which data
are available, thus allowing only a small laboratory
testing program with anticipated materials in the
appropriate planetary environment simulation. The
scraper edges for the dozer blade are often made of
rolled DH-2 steel [PC-11] for which data are also
available.

MISSION SCENARIOS

Unmanned. From Earth, launch the machine depicted
in Figure 3 attached to a lunar lander. Construct the
base as shown in Figure 7. Perform automated shutdown.

From consideration of the conceptual design of the
machine, three events which will result in complete
mission failure immediately come to mind:

(1) Primary power system failure,

(2) Excess sinkage (i.e. getting stuck due to
high ground pressure wheels, and

(3) Rollover/hangup.

Manned. Prepare vehicle for launch to Mars for
initial base construction there (e.g. replace worn
parts, insert Mars specific modules). Inspect, test,
and evaluate melt tracks, bricks, and compacted
materials. Place bricks. Test melt-tracks with a

rover type vehicle.

CONCLUSION

The use of rock melting and hot pressing
techniques for making building materials seems the most
appropriate approach at this time. The use of typical
adhesives such as portland cement and mortar is
considered to be impractical unless the heat intensive
methods outlined in this paper fail to produce useable
materials. Experimental data documented in the
literature on materials similar to materials proposed
as construction products indicate that useable
materials can be successfully produced.

With the present and near term future developments
in thermoelectric power generation and electric motors,
it is apparently feasible to manufacture a device which
can make planetary surface transportation systems and
protective structures. Considerable research into the
engineering properties of product materials is needed
before detailed design of the machine can be
accomplished. However, if the basic missions of the
machine outlined herein are considered appropriate, a
modular, conceptual design of the machine may be
performed which will minimize the effect of changing
technologies.

APPENDIX 4-A: REFERENCES and ADDITIONAL READING

1. Allen C.C., Gooding J.L., and Keil K.,
    "Hydrothermally Altered Impact Melt Rock and Breccia:
    Contributions to the Soil of Mars", Journal of
    Geophysical Research, Vol. 87, No. B12, 30 Nov.
    1982, pp. 10083-10101.

2. Altseimer J.H., "Technical and Cost Analysis of
    Rock-Melting Systems for Producing Geothermal Wells",
    LA-6555-MS, LASL, Nov. 1976.

3. American Ceramic Society, Inc., "Ceramic Source
    1986", Vol. 1, ACerS, 1985.

4. Armstrong P.E., "Subterrene Electrical Heater
    Design and Morphology", LA-5211-MS, LASL, Feb. 1974.

5. Balovnev V.I., "New Methods for Calculating
    Resistance to Cutting of Soil", USDA and NSF
    translation from Russian, Amerind Publishing Co., New
    Delhi, 1983.

6. Bazant Z.P., "Size Effect in Blunt Fracture:
    Concrete, Rock, Metal", Journal of Engineering
    Mechanics, ASCE, Vol. 110, No. 4, Apr. 1984, pp.
    518-535.

7. Bekker M.G., "Introduction to Terrain-Vehicle
    Systems", University of Michigan Press, Ann Arbor,
    1969.

8. Bennett G.L., Lombardo J.J., Rock B.J., "Nuclear
    Electric Power for Space Systems: Technology
    Background and Flight Systems Program", Intersociety
    Energy Conversion Engineering Conference, 16th,
    Atlanta, Ga., Aug. 1981, Proceedings, Vol. 1, ASME,
    N.Y., 1981, pp. 362-368.

9. Berry L.G., Mason B., and Dietrich R.V.,
    "Mineralogy", W.H. Freeman, N.Y., 1983.

10. Billingham J., Gilbreath W., and O'Leary B., eds.,
    "Space Resources and Space Settlements", NASA SP-428,
    NASA Scientific and Technical Information Branch,
    Washington, D.C., 1979.

11. Bradt R.C., Hasselman D.P.H., Lange F.F., eds.,

"Fracture Mechanics of Ceramics", Vol. 1-6, Plenum, N.Y., 1974-1983.

12. Brady G.S., and Clauser H.R., "Materials Handbook", McGraw-Hill, N.Y., 1986

13. Burmister D.M., "The Theory of Stresses and Displacements in Layered Systems and Applications to the Design of Airport Runways", Proceedings of the 23rd annual meeting, Highway Research Board, 1943, pp. 126-148.

14. Burmister D.M., "The General Theory of Stresses and Displacements in Layered Systems", Journal of Applied Physics, Vol. 16, No. 2, Feb. 1945, pp. 89-94.

15. Carr M.H., Saunders R.S., Strom R.G., Wilhelms D.E., "The Geology of the Terrestrial Planets", NASA SP-469, NASA, 1984.

16. Carrier W. III, Mitchell J.K., and Mahmood A., "The Nature of Lunar Soil", Journal of the Soil Mechanics and Foundations Division, ASCE, Oct. 1973, pp. 813-832.

17. Carrier W.D. III, Bromwell L.G., and Martin R.T., "Behavior of Returned Lunar Soil in Vacuum", Journal of the Soil Mechanics and Foundations Division, ASCE, Nov. 1973, pp. 979-996.

18. Caterpillar Tractor Co., "Caterpillar Performance Handbook", Peoria, Illinois, Oct. 1985.

19. Chesterman C.W., "The Audubon Society Field Guide to North American Rocks and Minerals", Alfred A. Knopf, N.Y., 1978.

20. Cook R.D., "Concepts and Applications of Finite Element Analysis", J. Wiley, N.Y., 1981.

21. Cort G.E., "Rock Heat-Loss Shape Factors for Subterrene Penetrators", LA-5435-MS, LASL, Oct. 1973.

22. Costes N.C., Farmer J.E., and George E.B., "Mobility Performance of the Lunar Roving Vehicle: Terrestrial Studies - Apollo 15 Results", NASA TR

R-401, Dec. 1972.

23. Criswell D.R., "Extraterrestrial Materials Processing and Construction", NSR 09-051-001 Mod. No. 24, Lunar and Planetary Institute, 30 Sep. 1978.

24. Crockford W.W., "Tensile Fracture and Fatigue of Cement Stabilized Soil", PhD dissertation, Texas A&M University, May, 1986.

25. Davies D.W., "The Relative Humidity of Mars' Atmosphere", Journal of Geophysical Research, Vol. 84, No. B14, 30 Dec. 1979, pp. 8335-8340.

26. Eshbach O.W., "Handbook of Engineering Fundamentals", J. Wiley, N.Y., 1966.

27. Fanale F.P., and Cannon W.A., "Mars: CO_2 Adsorption and Capillary Condensation on Clays-Significance for Volatile Storage and Atmospheric History", Journal of Geophysical Research, Vol. 84, No. B14, 30 Dec. 1979, pp. 8404-8414.

28. Freiman S.W., ed., "Fracture Mechanics Applied to Brittle Materials", ASTM STP 678, 1979.

29. Freiman S.W., and Fuller E.R., Jr., eds., "Fracture Mechanics for Ceramics, Rock, and Concrete", ASTM STP 745, 1981.

30. Gdoutos E.E., "Problems of Mixed Mode Crack Propagation", Martinus Nijhoff, The Hague, 1984.

31. Gopalaratnam V.S., Shah S.P., "Softening Response of Plain Concrete in Direct Tension", ACI Journal, May-Jun 1985, pp. 310-323.

32. Hanold R.J., "Viscous Melt Flow and Thermal Energy Transfer for a Rock-melting Penetrator: Lithothermodynamics", LA-UR-74-328, LASL, 1973.

33. Hanold R.J., "Large Subterrene Rock-Melting Tunnel Excavation Systems. A Preliminary Study.", LA-5210-MS, LASL, Feb. 1973.

34. Hanold R.J., "Rapid Excavation by Rock Melting --

LASL Subterrene Program -- December 31,
1972-September 1, 1973, LA-5459-SR, LASL, Nov. 1973.

35. Hanold R.J., "Rapid Excavation by Rock Melting --
    LASL Subterrene Program -- September 1973-June 1976",
    LA-5979-SR, LASL, Feb. 1977.

36. Hertzberg R.W., "Deformation and Fracture
    Mechanics of Engineering Materials", J. Wiley and
    Sons, N.Y., 1983.

37. Houston Chamber of Commerce, "Proceedings of the
    1974 Technology Transfer Conference", Houston, Tx.,
    1974.

38. Johnson S.W., and Leonard R.S., "Lunar-Based
    Platforms for an Astronomical Observatory",
    Proceedings of SPIE The International Society for
    Optical Engineering, Vol. '493, The National
    Symposium and Workshop on Optical Platforms, Wyman
    C.L. and Poulsen P.D., eds., Huntsville, Ala., Jun.
    1984.

39. Jones J.T., Berard M.F., "Ceramics, Industrial
    Processing and Testing", Iowa State University Press,
    1972.

40. Kachanov M., "Continuum Model of Medium with
    Cracks", Journal of the Engineering Mechanics
    Division, ASCE, Vol. 106, Oct. 1980, pp.
    1039-1051.

41. Kassir M.K., and Sih G.C., "Mechanics of Fracture
    2, Three-Dimensional Crack Problems", Noordhoff,
    Leyden, the Netherlands, 1975.

42. Koerner R.M., "Effect of Particle Characteristics
    on Soil Strength", Journal of the Soil Mechanics and
    Foundation Division, ASCE, Jul. 1970, pp.
    1221-1234.

43. Krupka M.C., "Selected Physicochemical Properties
    of Basaltic Rocks, Liquids, and Glasses", LA-5540-MS,
    Informal Report, Los Alamos Scientific Laboratory
    (LASL), Los Alamos, N.M., Mar. 1974.

44. Krupka M.C., "Thermodynamic Stability
    Considerations in the Mo-BN-C System. Application to

Prototype Subterrene Penetrators", LA-4959-MS, LASL, May 1972.

45. Mase G.E., "Schaum's Outline on Theory and Problems of Continuum Mechanics", McGraw-Hill, N.Y., 1970.

46. Mendell W.W., ed., "Lunar Bases and Space Activities of the 21st Century", Lunar and Planetary Institute, Houston, 1985.

47. Mindess S., and Young J.F., "Concrete", Prentice-Hall, 1981.

48. Mitchell J.K., Drozd K., Goodman R.E., Heuze F.E., Houston W.N., Willis D.R., and Witherspoon P.A., "Lunar Surface Engineering Properties Experiment Definition, Summary Technical Report", NASA contract NAS 8-21432, University of California Berkeley, Jan. 1970.

49. Mitchell J.K., Goodman R.E., Hurlbut F.C., Houston W.N., Willis D.R., Witherspoon P.A., and Hovland H.J., "Lunar Surface Engineering Properties Experiment Definition, Summary Technical Report", NASA contract NAS 8-21432, University of California Berkeley, Jul. 1971.

50. Mitchell J.K., Houston W.N., Scott R.F., Costes N.C., Carrier W.D. III, and Bromwell L.G., "Mechanical Properties of Lunar Soil: Density, Porosity, Cohesion, and Angle of Internal Friction", Proceedings of the Third Lunar Science Conference, Supplement 3, Geochimica et Cosmochimica Acta, Vol. 3, M.I.T. Press, 1972, pp. 3235-3253.

51. Mitchell J.K., Houston W.N., Carrier W.D. III, and Costes N.C., "Apollo Soil Mechanics Experiment S-200 Final Report Covering work performed Under NASA Contract NAS9-11266", Space Sciences Laboratory Series 15, Issue 7, Geotechnical Engineering, University of California, Berkeley, Jan 1974.

52. Mizutani H., Spetzler H., Getting I., Martin R.J. III, Soga N., "The Effect of Outgassing Upon the Closure of Cracks and the Strength of Lunar Analogues", Proceedings of the Eighth Lunar Science Conference, Supplement 8, Geochimica et Cosmochimica

Acta, Vol. 1, Pergamon, 1977, pp. 1235-1248.

53. Moore H.J., Hutton R.E., Scott R.F., Spitzer C.R.,
and Shorthill R.W., "Surface Materials of the Viking
Landing Sites", Journal of Geophysical Research, Vol.
77, No. 28, 30 Sep 1977, pp. 4497-4523.

54. Moore H.J., Clow G.D., Hutton R.E., "A Summary of
Viking Sample-Trench Analyses for Angles of Internal
Friction and Cohesions", Journal of Geophysical
Research, Vol. 87, No. B12, 30 Nov. 1982, pp.
10043-10050.

55. Nielsen R.R., Abou-Sayed A., and Jones A.H.,
"Characterization of Rock-Glass Formed by the LASL
Subterrene in Bandelier Tuff", Terra Tek Inc., Salt
Lake City, Utah, TR 75-61, Nov. 1975.

56. Pariseau W.G., "Gravity Flow of Powder in a Lunar
Environment", Report of Investigations 7577, U.S.
Department of the Interior, Bureau of Mines.

57. Rowley J.C., "Potential for Tunneling Based on
Rock and Soil Melting", LA-UR, LASL, Apr. 1985.

58. Ryder G., and Norman M.D., "Catalog of Apollo 16
Rocks", NASA Curatorial Branch Publication 52, Sep.
1980.

59. Sih G.C., ed., "Mechanics of Fracture 3, Plates
and Shells with Cracks", Noordhoff, Leyden, the
Netherlands, 1977.

60. Simiu E., and Scanlan R.H., "Wind Effects on
Structures: An Introduction to Wind Engineering", J.
Wiley and Sons, N.Y., 1978.

61. Soga N., Spetzler H., and Mizutani H., "Comparison
of Single Crack Propagation in Lunar Analogue Glass
and the Failure Strength of Rocks", Proceedings of
the Tenth Lunar and Planetary Science Conference,
Supplement 11, Journal of the Geochemical Society and
the Meteoritical Society, Vol. 3, Pergamon, 1979,
pp. 2165-2173.

62. Spangler M.G., and Handy R.L., "Soil Engineering",
Harper and Row, N.Y., 1982

63. Stanton A.E., "Heat Transfer and Thermal Treatment Processes in Subterrene-Produced Glass Hole Linings", LA-5502-MS, LASL, Feb. 1974.

64. Timoshenko S.P., and Goodier J.N., "Theory of Elasticity", McGraw-Hill, N.Y., 1970.

65. Tittmann B.R., Clark V.A., and Spencer T.W., "Compressive Strength, Seismic Q, and Elastic Modulus", Proceedings of the Eleventh Lunar and Planetary Science Conference, Supplement 14, Journal of the Geochemical Society and the Meteoritical Society, Vol. 3, Pergamon, 1980, pp. 1815-1823.

66. Ugural A.C., "Stresses in Plates and Shells", McGraw-Hill, N.Y., 1981.

67. Van Vlack L.H., "Materials for Engineering: Concepts and Applications", Addison-Wesley, Mass., 1982.

68. Westergaard H.M., "New Formulas for Stresses in Concrete Pavements of Airfields", Transactions, ASCE, Vol. 113, 1948, pp. 425-444.

69. Williams R.J., and Jadwick J.J., "Handbook of Lunar Materials", NASA Reference Publication 1057, 1980.

70. Wittman F.H., ed., "Fracture Mechanics of Concrete, Developments in Civil Engineering", 7, Elsevier, N.Y., 1983.

71. Yoder E.J., and Witczak M.W., "Principles of Pavement Design", J. Wiley and Sons, N.Y., 1975.

PERSONAL CONTACT REFERENCES

PC-1. Alred J., NASA/ED2, (713) 483-4478.

PC-2. Blanchard D.P., NASA/SN2, (713) 483-3274.

PC-3. Brennan T., Spectra-Physics, (713) 688-8814.

PC-4. Bufkin A., NASA/ED2, (713) 483-2536.

PC-5. Coker B., Tom Fairey Company (John Deere), (713) 695-9770.

PC-6. Fox M., San Jacinto College, (713) 476-1501 (x527)

PC-7. Gee J.E., Caterpillar Tractor Co., Peoria Proving Ground, (309) 698-5959.

PC-8. Harvey H.W., Ingersoll-Rand Compaction Division, (717) 532-9181.

PC-9. Heller C., Sr., WED Transportation Systems, Inc., Subsidiary of Walt Disney Co., (713) 821-0121.

PC-10. Jansen L.T., Sundstrand Mobile Controls Applications, (612) 559-2121.

PC-11. Kliment S.C., Mustang Tractor and Equipment (Caterpillar), (713) 460-2000.

PC-12. McCormick J.A., Ingersoll-Rand Corporate Advertising, (201) 689-4554

PC-13. Meek T.T., LASL, Materials Science and Technology Division, (505) 667-2129.

PC-14. Pope D., Spectra-Physics, (800) 538-7800, (not contacted).

PC-15. Rowley J.C., LASL, Earth and Space Sciences Division, (505) 667-1378.

APPENDIX 4-B: LIST OF SYMBOLS

a	radius
b	constant
c	cohesion
E	elastic stiffness (Young's modulus if isotropic)
g	earth gravitational acceleration
g_m	Mars gravity
g_o	Lunar gravity
JIc	Fracture toughness, Energetic
k	thermal conductivity
k_c	constant in Bekker model
KIc	Fracture toughness, Stress intensity factor
ksg	modulus of subgrade reaction
k_ϕ	constant in Bekker model
N	sample size
n	exponent in Bekker model
q	distributed load
r	radius
SG	specific gravity
T	Temperature
z	depth
α	linear coefficient of thermal expansion
β	angle of repose
Θ	angle
ν	Poisson's ratio
ρ	density (used interchangeably as unit weight)
σ_c	compressive strength
σ_y	yield strength (or indirect tension or ultimate tensile strength)
ϕ	angle of internal friction

AYER	LASL system software
DDM	NASA/JSC system software
IDEAS**2	NASA/JSC system software
JSC	Johnson Space Center
LASL	Los Alamos Scientific Laboratory
LRV	Lunar Roving Vehicle
MSFC	Marshall Space Flight Center
NASA	National Aeronautics and Space Administration

APPENDIX 4-C: DATA FROM MANUFACTURER'S LITERATURE

Manufacturer	Model	HP (kW)	Operating Weight (kg)	Approximate Machine Volume (m^3)	Dozer/Loadr Available, Capacity (m^3)
Caterpillar	CB214	24.0	2300	5.4	N
Caterpillar	CB224	24.0	2450	6.4	N
Caterpillar	CB314	41.0	3357	6.7	N
Caterpillar	D3B	48.0	6915	11.2	1.21
Caterpillar	IT12	48.0	7554		Y
Caterpillar	CP323	52.0	4560	20.8	Y
Caterpillar	CB414	52.0	5780	14.8	N
Caterpillar	CS431	52.0	6110	20.5	N
Caterpillar	D3BC75	56.0	7371	11.2	1.21
Caterpillar	D4E	60.0	9090	13.2	1.7
Caterpillar	CP433	60.0	6750	23.6	Y
Caterpillar	CS433	60.0	6720	27.2	N
Caterpillar	IT18	63.0	8660		Y
Caterpillar	D4H	67.0	12252	25.0	1.89
Caterpillar	PR105	67.0	7711	26.3	N
Caterpillar	CB514	68.0	9730	22.5	N
Caterpillar	D5B	78.0	11619	21.2	2.57
Caterpillar	IT28	78.0	9633		Y
Caterpillar	D5H	90.0	13890	32.4	2.66
Caterpillar	518	96.0	14243	60.3	Y
Caterpillar	D6D	104.0	15695	23.6	3.06
Caterpillar	CB614	115.0	11340	34.4	N
Caterpillar	CS551	115.0	10400	41.3	N
Caterpillar	CP553	116.0	12200	50.4	Y
Caterpillar	CS553	116.0	10780	43.0	N
Caterpillar	D6H	123.0	20612	40.0	4.08
Caterpillar	D7G	149.0	20666	30.6	6.42
Caterpillar	814B	157.0	20580	54.8	2.91
Caterpillar	815B	157.0	20037	87.5	Y
Caterpillar	816B	157.0	20628	86.0	Y
Caterpillar	D7H	160.0	22796	46.1	6.42
Caterpillar	824C	231.0	30380	72.9	4.67
Caterpillar	826C	231.0	31310	115.5	Y
Caterpillar	D8L	250.0	37417	66.6	13.6
Caterpillar	834B	336.0	46355	98.1	7.27
Caterpillar	D9L	343.0	52478	71.6	18.5
Caterpillar	D10	522.0	79619	95.1	29.07
Dresser	VOST2-42,	25.7	3257	12.1	N
Ingersoll-Rand	DA30	24.0	3200	7.2	N
Ingersoll-Rand	DA28	24.6	2275	4.7	N
Ingersoll-Rand	DA40	57.0	6990	13.5	N
Ingersoll-Rand	SPA50	70.0	7410	24.4	N
Ingersoll-Rand	SPA56	70.0	9160	31.3	N
Ingersoll-Rand	SP48	70.0	6600	20.1	N
Ingersoll-Rand	SP48DD	70.0	6920	23.0	N

|					
Ingersoll-Rand	SPF48	70.0	7240	24.2	N
Ingersoll-Rand	SP56	70.0	8913	29.5	N
Ingersoll-Rand	SD100	71.6	10500	29.5	N
Ingersoll-Rand	SD100F	82.8	11500	29.5	N
Ingersoll-Rand	SP56DD	84.0	9389	33.0	N
Ingersoll-Rand	SPF56	84.0	10206	33.5	Y
Ingersoll-Rand	DA48	86.0	9099	21.1	N
Ingersoll-Rand	DA50	86.0	10020	28.2	N
Ingersoll-Rand	SP84	123.0	13900	36.2	N
Ingersoll-Rand	SPF84	123.0	14210	36.2	N
Ingersoll-Rand	DS84	146.0	20684	34.7	N
Ingersoll-Rand	DF84	146.0	18733	43.3	N
Ingersoll-Rand	LF450	160.0	20455	100.2	Y
Ingersoll-Rand	SP60DD	164.0	17101	45.2	N
Ingersoll-Rand	SPF60	164.0	18870	46.4	N
Ingersoll-Rand	SPF60C	164.0	20140	49.6	N
Ingersoll-Rand	LF750	302.0	35835	148.5	Y
John Deere	655	11.9	732	5.3	Y
John Deere	316	11.9	354	2.1	Y
John Deere	650	12.0	744	5.5	Y
John Deere	330	12.0	408	2.1	Y
John Deere	318	13.4	354	2.1	Y
John Deere	755	14.9	764	5.9	Y
John Deere	430	14.9	533	3.9	Y
John Deere	750	15.0	907	5.1	Y
John Deere	855	17.9	827	4.2	Y
John Deere	30	18.0	3000		N
John Deere	850	19.0	1222	7.0	Y
John Deere	570	23.0	1515	6.9	Y
John Deere	50	29.0	4355		N
John Deere	675	32.8	1991	9.2	Y
John Deere	355D	36.0	5625	14.2	0.75
John Deere	350D/630S	36.0	4810	18.7	Y
John Deere	350D	36.0	5465	20.9	Y
John Deere	70	41.0	6620		N
John Deere	444D	67.0	9595	38.7	1.15
John Deere	544D	86.0	10820	47.9	1.34
John Deere	655B	90.0	15240		2
John Deere	750B/650S	90.0	14900	53.0	Y
John Deere	750B	90.0	13489	39.0	Y
John Deere	750BLGP	90.0	15806	55.0	Y
John Deere	755B	104.0	17000		2.25
John Deere	850B	123.0	20124	66.9	Y
Boeing	LRV		707	7.3	N
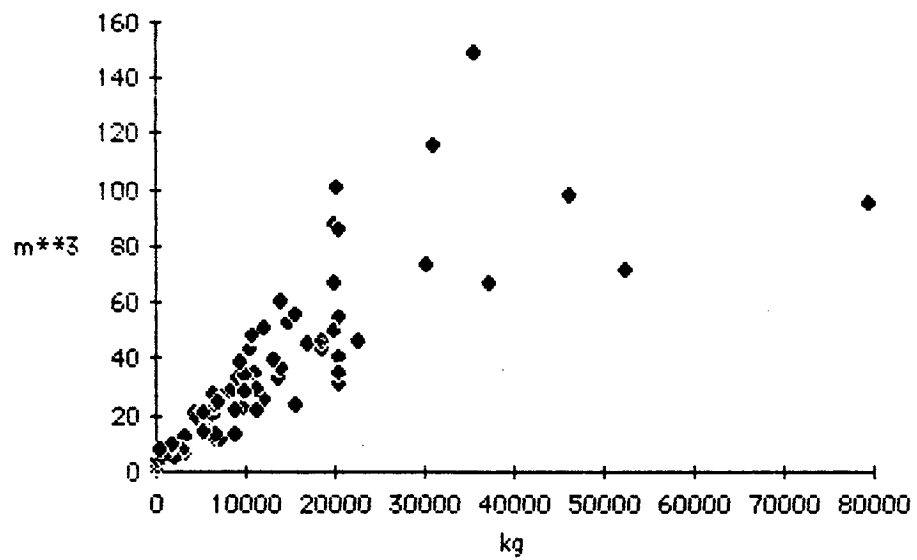
Figure C-1. Volume versus weight for various construction
machinery.

APPENDIX 4-D: SELECTED ASTM TESTS *

C7   Paving brick
or
C131   Resistance to abrasion of small size aggregate
          by use of the Los Angeles machine
or
C418   Abrasion resistance of concrete by sandblasting
or
C779   Abrasion resistance of horizontal concrete
          surfaces

C67   Sampling and testing brick and structural clay
         tile

C158   Flexure testing of glass

C589   Apparent impact strength of preformed
         block-type insulating materials
or
E23   Notched bar impact testing of metallic materials

C598   Annealing point and strain point of glass by
         beam bending

C623   Young's modulus, shear modulus, and Poisson's
         ratio for glass and glass-ceramics by resonance

C637   Aggregates for radiation-shielding concretes

C638   Constituents of aggregates for
         radiation-shielding concrete

D1195   Repetitive static plate load tests of soils and
         flexible pavement components, for use in
         evaluation and design of airport and highway
         pavements

D1883   Bearing ratio of laboratory-compacted soils

D3397   Triaxial classification of base materials,
         soils, and soil mixtures

D4535   Measurement of thermal expansion of rock
         using a dilatometer

E18   Rockwell hardness and Rockwell superficial

hardness of metallic materials

E399   Plane-strain fracture toughness of metallic
        materials

E510   Determining pavement surface frictional and
        polishing characteristics using a small
        torque device

E647   (on fatigue crack propagation)

E813   JIc, a measure of fracture toughness

*   Compiled primarily from the 1979 annual books
of ASTM standards. Applicable new standards and
modifications to the above standards may exist.

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

Texas A&M University

Digital Data From Shuttle Photography:
The Effects of Platform Variables

Prepared by: Bruce E. Davis, Ph.D.

Academic Rank: Assistant Professor

University & Department: Univ. of Hawaii, Hilo

                                      Department of Geography

NASA/JSC

Directorate: Space and Life Sciences

Division: Planetary Exploration

Branch: Experimental Planetology

JSC Colleague: Dr. David Amsbury

Date: August 15, 1986

Contract #:   NGT-44-005-803

ABSTRACT

Two major criticisms of using Shuttle hand-held photography
as an Earth science sensor are that it is non-digital/
non-quantitative and that it has inconsistent platform
characteristics, e.g., variable look-angles, especially as
compared to remote sensing satellites such as Landsat and
SPOT. However, these criticisms are assumptions and have
not been systematically investigated. This project
focusses on the spectral effects of off-nadir views of
hand-held photography from the Shuttle and their role in
interpretation of lava flow morphology on the island of
Hawaii. The first stages of the research are
discussed--digitization of photography at JSC and use of
LIPS image analysis software in obtaining data.
Preliminary interpretative results of one flow are given.
Most of the time was spent in developing procedures and
overcoming equipment problems. Preliminary data are
satisfactory for detailed analysis.

NASA Colleague: Dr. David Amsbury SN4 x6451

Since 1972, Landsat has provided the standard for Earth science remote sensing from space, while hand-held photography from manned spacecraft has been relegated to an inferior subsidiary role of illustration. Despite the provision of some 37,000 hand-held photographs from the Space Transportation System Shuttle program in the past five years, there have been very few scientific reports using the imagery as a major data source (see e.g., Kaltenbach, Helfert, and Wells, 1984; Gallegos, et al., 1984; Nerem and Holz, 1984). The fundamental reason for diversity of scientific acceptance between the two sensors is a perception of technological performance differences: Landsat is seen as state-of-the-art but photography is deemed deficient and largely second-rate. For example, whereas Landsat is a digital scanner system which renders data computer compatible, photography is film based--visually analog and not intrinsically amenable to computer analysis techniques. Hand-held photography (herein referred to as HHP) also has variable platform characteristics, such as a diversity of viewing-angles (also called look angles and usually referred to as off-nadir when non-vertical angles are considered) and different solar illumination conditions. Landsat, to its supposed advantage, is very consistent in platform attributes (especially the nadir view), thereby offering a

standard format of imagery. Believing that HHP can be a more useful scientific data resource, the differences in platform performance were called to questions with the hypothesis that non-vertical viewing angles and attendant illumination changes are not debilitating in terms of image quality and utility, and in fact, may be advantageous in some earth science research. This study investigates the nature of off-nadir viewing from HHP and its role in analysis of lava flow morphology on the island of Hawaii.

The basic objective of this project was to develope a methodology to analyze spectral (tonal) differences of given features on HHP frames having variable look angles. Three phases of the project were perceived: digitization of photography (the processing stage), development of data extraction techniques, and analysis of selected features (applications stage). Most of the summer period involved the first two steps and preliminary analysis of data was initiated. This report discusses the theoretical and practical considerations of digitally processing photography and the means to extract useful information. Preliminary results of one research site are presented as a demonstration of the methodology.

THEORY

A Lambertian surface is a perfectly diffusing target that
reflects energy equally in all directions (e.g., Slater,
1983). The physics of such perfection are calculable but
terrestrial surfaces are not Lambertian, so there is
interest in the interactions of electromagnetic energy with
various types and conditions of targets. Off-nadir sensing
has been investigated, but primarily for vegetated surfaces
(e.g., Bartlett, Johnson, Hardisky, Klemas, 1986; Daughtry
and Ranson, 1986; Gerstl and Simmer, 1986; Goel and
Deering, 1985; Holben, Kimes, Fraser, 1986; Kimes, Newcomb,
Nelson, Schutt, 1986; Li and Strahler, 1985; Lord,
Desjardines, Dubé, Brach, 1985; Norman, Welles, Walter,
1985; Simmer and Gerstl, 1985). No research on the effects
of different look angles to spectral response of lava flow
(or any volcanic landscapes) could be found. Further,
emphasis has been on the derivation of mathematical models
of off-nadir viewing but this investigation concentrates
initially on detection of changes on a new medium and their
role in analysis of lava flow morphology.

Basic theory of digital imagery processing and analysis
forms the framework of this project (see Jensen, 1986, for
the best single reference) and details of its application,

i.e., specific procedures, are discussed in the next section. The literature on digital remote sensing is vast but will be mentioned only when directly relevant; no literature review is offered here. As prescribed above, digitization is considered first, followed by data extraction and data analysis.

Photography, a film-based image, is a visual product and analog by nature. Conversion to a digital format, necessary for computer analysis, is a theoretically simple process of re-imaging with a vidicon camera (in this case, but a scanning densitometer also can be used) and transformation of picture elements (pixels) into discrete values representing tone or density. Specifically, light of a known intensity is transmitted through a transparency and the reduction in intensity received by the vidicon, on a pixel by pixel basis, is translated as film density. Pixels, having specific X-Y locations, are restructured into a digital image of the photographed scene. Color photographs are reduced to three distinct bands (usually red, green and blue) by the use of filters in the re-imaging process, making a three-band digital image, ready for computer manipulation.

Digitization of photography is not a new technique (e.g., Hoffer, Anuta, Phillips, 1971; Jensen, Estes, Tinney, 1978;

LeSchack, 1971; McDowell, 1974; Scarpace, Quirk, Kiefer, Wynn, 1981; Smedes, Linnerud, Woolaver, Hawks, 1971), but it is not a popular technique for quantitative analysis, primarily due to the many vagaries inherent in photography and difficulty in control and repeatibility of the digitizing process. Despite high quaility control, the chemical processes in manufacturing, storing, development, and reproduction of films are not quantitatively known or maintained, and thereby are subject to latitudes of variance. Because of these factors, precise analytical controls are very difficult to attain and derivation of absolute spectral signatures, for example, is seldom attempted. In fact, Scarpace (1978, p. 1287) maintains that "film density is not the parameter to be correlated with the reflected energy from the ground. The dye densities formed in films depend in a non-linear way on not only the amount of energy and its spectral distribution striking the unexposed film, but also on the processing of the imagery." Radiometric calibration derives the relationship between film density and light energy (making D-log E curves of exposure and film response) but the process is too complex to include here (involving photo technicians, and detailed dye measurements) (also see McDowell and Specht, 1974).

The result of these difficulties is that while digitization

is a relatively simple process, as described, great care must be taken in maintaining data that are comparative. A major aim of this project (and most of the work expended) was in development of procedures for producing quantitatively standard data so that features on different frames could be compared effectively and accurately.

There is no practical way to determine film chemistry processes and their quantitative meaning for HHP. The best approach is to use reproductions as close as possible to first generation images (keeping data degradation to a minimum) and to select frames on the same roll (to ensure reasonably identical processing). Fortunately, two four-frame sequential series of second generation images of the desired study area were available, minimizing extraneous factors of variability.

The initial aim of data analysis was to spectrally characterize lava flows at various viewing angles and to evaluate the signature differences.[1] Spectral signatures of changing linear features are best represented as a chain

----

1. As noted, absolute spectral signatures are not attempted here and only relative responses are used. The term "spectral signature" is used for convenience and does not imply establishment of pure signatures; actually, the proper term to use would be film response.

of tonal values along their paths. That is, pixel values from each band are collected at locations along the lava flows and compiled into sets of value curves. Various statistical methods are available that determine differences and relationships of curves and points within a set of data. Analysis of single frame information may be useful in terms of lava flow morphology, but when all frames are combined and compared, synergetic interpretation may be possible.

One of the major problems of multi-frame image analysis is obtaining data that are truly and purely comparative. Under perfect (unrealistic) conditions, all pixel values reflect ground variables only and photographic or digitizing factors do not induce tonal corruption. However, practical conditions contribute both noise and potentially invalid values. For example, Jensen (1986, p. 16) states that vignetting is one of the most serious problems in video digitizing--fall-off of intensity from the center of output to the edges, usually in a circular pattern. Also, Jensen, Estes, Tinney (1978) found vignetting to be a critical component preventing consistent measurements of agricultural fields.

If such intrinsic "antagonists" can be identified and measured, their dilutions can be subtracted from data. The

aim, then, is to "normalize" data to some common base or measure so that only the ground effects contribute to tonal change. From there, changes wrought by different look angles and slant ranges can be deduced. There are numerous methods of normalization, but a fairly simple approach was used that identifies and compensates for extraneous influence, i.e., light table variances were identified and removed. Once extrinsic factors have been eliminated and a satisfactory set of densities values gathered, analysis of data can proceed.

DATA

As stated, two ideal sequences of images of Hawaii were located: visual 65A-50-056 to 059 and color infrared 65A-55-105 to 108. Table 1 presents their viewing geometries of the visual color frames. Initially, three lava flows were examined--Keamuku (about 300 years old), 1859, and 1880 (Map 1). To maximize resolution and spectral information, enlarged subimages of each flow were produced from each of the frames above. Further, in that each band of the images requires individual analysis, there was a total of 72 data images (8 frames x 3 bands each x 3 study flows).

The photographs were digitized on JSC's new Gould Eikonix

TABLE 1: IMAGE GEOMETRY

Based on Shuttle to Keamuku flow
Altitude: 176 n.m.

Frame	Ground Range (n.m.)	Slant Range (n.m.)	Look Angle (degrees)	Solar[1] Az	Elev
56	315	372	59	141.0	53.0
57	229	293	51	143.6	52.5
58	210	278	49	144.2	52.3
59	160	241	42	146.1	51.8

1. Solar Az: Solar Azimuth from Shuttle to Sun, measured cw from north.
   Solar Elev: Solar elevation angle at Shuttle nadir.
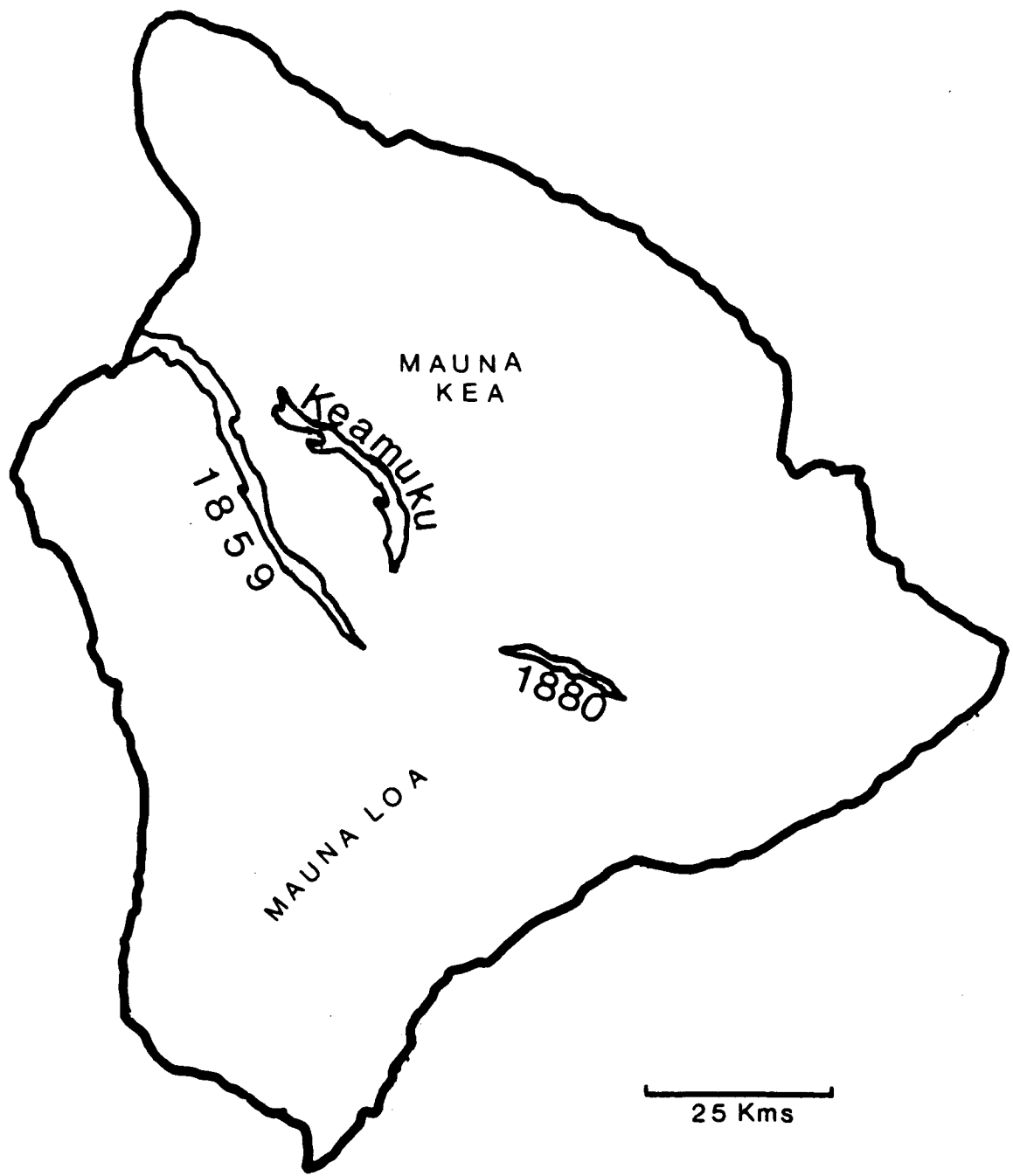
MAP 1 HAWAII

850 digitizing system in Building 17. It offers digitizing formats of 512 x 512 pixels, 1024 x 1024, and 2048 x 2048, each with center or upper left corner (0,0) origins. Additionally, either transparencies or prints may be used, though the former usually present best results. Computer processing and analysis were performed with LIPS--Library of Image Processing Software.

METHODS AND PROCEDURES

Indicated above were the contrasts between theory and practice in digitization and analysis of imagery. Discussed here are the specific procedures used (as guided by basic theory) and the problems encountered. Practical considerations are addressed as a means for developing insight and furthering project evolution. The format employed is representative of a digitizing task and subsequent preparation for image analysis. Specific LIPS commands and responses are given for JSC operator convenience.

Digitizing System:

Light Table: Consisting of floodlights for opaque prints and a small table window with lighting from underneath for transparencies (the format used here), the light table

provides illumination for the images. It is this component which created intransigent problems and seriously delayed progress. Illumination must be strong so that other variablesin the total system can be exercised. However, because of unsuspected power supply wiring problems, light intensity was a fraction of its design, therefore necessitating maximum conditions of other controls just to attain minimum illumination. Consequently, poor and unacceptable image data resulted, causing much waste in time and effort. The problem was finally rectified in week nine of the ten week program, so data are preliminary, incomplete, and essentially unanalyzed.

Another major (and continuing) problems of the lighting system is that it is a condenser type, resulting in uneven illumination across viewing area. A "hot spot" is particularly evident (though always present) at camera positions outside the narrow focus position--apparently inherent in condenser systems. A diffusing light system (more expensive) retards such conditions and offers more even lighting. Perhaps this also is inherent in Eikonix instrumentation, as other users have complained of similar limitations. The applied aspects of such problems is that standardization of pixel values is not ensured, forcing development of means to remove the variance. Further delays were encountered because of mathematical limitations

in LIPS (to be discussed). In essence, the illumination
component is a critical weakness to spectral analysis work
in this digitizing system.

Camera: the vidicon camera digitizes with the use of a
linear array of 2048 photoelectric detectors that scan the
image in a vertical manner. Exterior controls consist of
f-stop, focus, framing, and lens. Framing is accomplished
by vertical racking of the camera on its support, but as
mentioned, best results are attained at the condenser
system's focus point, which greatly limits framing and
scaling. The only reasonable option is to use different
lens--normally available are a 50mm, 105 mm, and extension
tube, although these too must be positioned at the lighting
system's focus. Also included in the camera is a filter
wheel which introduces red, green, and blue filters in the
light path (after the lens) in order to reduce a color
photo into three primary colors.

A major problem in the camera is the presence of numerous
dead or deficient detectors, which causes streaking and
erroneous data on the image. A column of four or five dead
detectors in the center creates a wide dark band that is
visually distracting and quantitatively distorting in data
analysis. For example, a pixel value histogram of a
central bright area will show values of 0 (dark), which

obviously do not belong and also create erroneous frequency statistics. Designing images that fit into the limited framing and that avoid most of the dead detectors is unnecessarily time consuming and can cause potentially undesirable compromises.

Control Box: Other major exterior controls for digitizing are contained on an instrument beside the light table. Light intensity is monitored with an oscilloscope atop the box and controlled with an incremented lever. Filters (or clear) are selectable, as are several other non-critical exposure variables that are usually controlled with computer inputs. Fine focussing is aided with a contrast reading on the oscilloscope.

Software: Primary regulation of digitizing is in the software, using IDTST, CALIBRATION and DIGIT programs. IDTST establishes detector exposure time (dwell time) and scan rate. CALIBRATION sets the range of detector response from the darkest input (to 0) to the brightest (to 255). DIGIT prepares the system for digitizing and within are selections for exposure and scan framing. Image size can be dictated by choosing one of three scan outputs (512 x 512 lines and pixels [columns], 1024 x 1024, or 2048 x 2048) and selecting a center or upper left origin. Once digitizing has been accomplished, entry into LIPS begins

the image manipulation and analysis phases.

The newly digitized scene is transferred from the primary operating system (VMS) to LIPS by a DEFIMG (DEFine IMaGe) routine. For scan outputs over 512 x 512, a screen cursor can be positioned anywhere on the scene to define a 512 x 512 subimage. Once the desired images have been saved, normal LIPS operations can begin, e.g., enhancements, algebraic renditions, signature analysis, etc.

Under ideal conditions, the process for signature analysis would include digitizing as described above, saving the image in LIPS, and using one of several pixel value collection routines--point value, box histogram, or profile, all manually controlled. Organization and statistical analysis of resulting data would then round out the task. The process is fairly straightforward and seemingly easy, but as indicated, numerous barriers were encountered, the first of which was inadequate light table intensity.

There simply is no solution to insufficient lighting for images. A normal image's wide range of values are compressed at the lower end of the brightness range and spectral resolution and contrast are lost. Contrast stretching offers visual relief but doesn't change the

relationships of original data. Therefore, satisfactory illumination is the first crucial criterion for progress.

While awaiting solution of the light intensity problem, the uneven illumination was investigated. A fundamental premise of this research is that pixel values in a scene represent ground features and that different values within a feature indicate inhomogeneity of surface appearance or atmospheric effect, which in turn stimulates investigation into the reasons for these differences. Hence, one has to be assured that pixel values are results of factors inherent in the feature or in the sensing (atmospheric conditions, solar angles, etc.) and are not the product of treatment of the image in the laboratory. Thus, distortion induced by image analysis equipment must be removed.

In the simplest sense, irregularities can be compensated for by adding or subtracting the difference between observed and the desired. However, care must be taken in that simple addition or subtraction can change the relationship of values, particularly when two different exposure intensities are used in measuring. More precisely, the problem is that the bare light table's illumination is far brighter than when covered with an image--the photo reduces light intensity so much that identical camera or exposure setting are incompatible with

the level of the bare light table. So, to measure the initial intensity and accompanying irregularities, exposure must be reduced as compared to image digitizing intensity. Exposure differences present problems in standardizing pixel values because of the wholely changed nature of the illumination. By computing on a ratio basis, however, most of the difference and variance can be managed. Using the maximum intensity of the light table as the standard (to which variances must be increased to) Formula 1 was used to produce a "mask" for each band (red, green, and blue). This and other formulae are applied to each color band of the image. LIPS commands are given in brackets where appropriate.

Formula 1:
[RATIO]

$$Mxy = \frac{Hc}{Vxy} \times 100$$

Where

M = Mask
xy = Coordinates of a given pixel
H = Maximum intensity value in entire scene
V = Intensity value

The mask then could be applied to the image for "normalization" of each band using the idealized formula:

Formula 2:

$$N = \frac{(Mxy)(Ixy)}{100}$$

5-19

Where

        N = Normalized image
        M = Mask value
        I = Image

These formulae correct light table illumination variances but LIPS presents difficulties in their application. The first problem is minor: LIPS can output only integers. Understandably, pixel values must be integers and at the end of any calculation of two images, LIPS is prepared to depict results in the form of an image. Consequently, rounding occurs; LIPS has no way of knowing that Mask, for example, is an intervening calculation where image output is not necessary. Therefore, resulting pixel values are plus or minus one DN (density number), an acceptable error for this project.

Mathematically, these formulae can be reduced to:

Formula 3:
$$N = \frac{H(I_{xy})}{V_{xy}}$$

A second problem is that LIPS does not accept designed formulae but uses only established routines. As in Formula 1, RATIO can manage A/B but not (AxB)/C; thus, the numerator must be constructed appropriately. A third intrinsic problem is introduced: output is limited to the range 0-255. Again, because LIPS is prepared to produce an image from any algebraic routine and because the tonal

range is restricted, any calculation performed must result in integers between (and including) 0-255. Multiplying H x I surely will exceed this limit, so LIPS has scaling subroutines to keep results within its boundaries. The user, then, cannot have confidence that each pixel has the desired (or same) relationship with all other pixels. The simple illumination rectification formula cannot be applied directly.

With some tenacity the problem can be overcome. Using one of LIPS' multiplication routines the modified (scaled) product of two images can be produced. In this case, the mask (Formula 1) is multiplied by the image, i.e., the numerator operation in Formula 2. To find the scaling number (transformation number) approximately a dozen sample pixel values are taken, using a floating cursor point value routine (PIxel). If a particular class of features is under study, pixel values should be taken from them. These values are compared to hand-calculated values of the Formula 2--the desired output. Once the ratio of LIPS product to desired output is derived, its integer value (ratio percentage x 100) is stored as a single-value image (using the CONSTANT routine in LIPS). Finally, the normalized image is achieved by ratioing the product image and transformation number x 100. These steps are:

1.   Formula 4: $P_{xy} = (M_{xy})(I_{xy})$
[MUL8 or MULI]

Where
     P = Intermediate product

2.   Obtain pixel value of selected sites (V_{xy}) using PI in LIPS.

3.   Using sample pixel values, calculate output of the desired
formula (Formula 2).

4.   Calculate the Transformation number for each band (T) by:

Formula 5: $T = \dfrac{P_{xy}}{N_{xy}} \times 100$

[The N_{xy} here is from Formula 2]

5.   Store T as a single-value image.

6.   Normalize by:

Formula 6: $N = \dfrac{P_{xy}}{T}$

7. Save P_{xy} as the final, normalized image to use in signature
analysis. Combine the color bands of each frame into a
three-band color image.

These procedures can be applied to full frames, but the

features under study here (lava flows) are too small on

each frame to present sufficient size for pixel value

sampling. As discussed, subimages can be produced at the

DEFIMG stage, thereby making an enlarged image for each

flow.   A total of 72 images serves as the data base.

5-22

Unfortunately, the above procedure must be applied to each subimage for normalization. Although some short-cuts could be performed, there is no assurance that standardization of images will result, primarily because the variances in illumination is spatially unique and one site's of numbers will not work for another site. For example, the transformation numbers for frame 059 was applied unsuccessfully to 056. The method is time-consuming, but serves as the only reasonable technique for overcoming inherent problems in the Eikonix hardware and in LIPS.

DATA ANALYSIS

By the time the problems were solved, little time remained for generation of complete data. The Keamuku flow was subimaged and pixel values were taken. Procedures of digitizing, normalization, and preliminary data collection were verified. No analysis of consequence has been completed but demonstration results are given and discussed. Because this report concentrates primarily on procedures and because complete discussion of data and analysis normally is lengthly (site and feature characteristics, literature history, tables of data, details of analysis, consideration of results, etc.), only brief remarks of results and preliminary thoughts of meaning (no conclusions) are presented.

Contrast normally decreases with increased viewing angle and with decreasing solar angle. Distance is not a major determinant but the increased atmosphere between sensor and target also tends to reduce contrast (e.g., Slater, 1983). Detail and features within lava flows are naturally very low contrast and the effects of look angle and solar illumination may be critical in lava morphology interpretation. The first stage of data analysis compares the spectral effects of the visual frames (056 to 059).

Figure 1 shows tonal response (DN) of the three color of frame 59 (the highest viewing angle--least off-nadir) to distance from the main vent. Although there is practically no vegetation (lichen) on the Keamuku flow, green has the brightest response. A comparison with frame 56 (lowest viewing angle) in Figure 2 shows the effect of changing look angle. Green retains the highest values in 56, but the difference between blue and red is more pronounced. Table 2 presents a selected set of correlations for further and quantitative comparison. All colors of frame 56 have good correlations with distance, as does red and green in 59, but blue in the latter image has almost no correlation. This is surprising given that blue should be scattered much more on 56 because of the greater viewing angle and slant range atmosphere and thus should exhibit
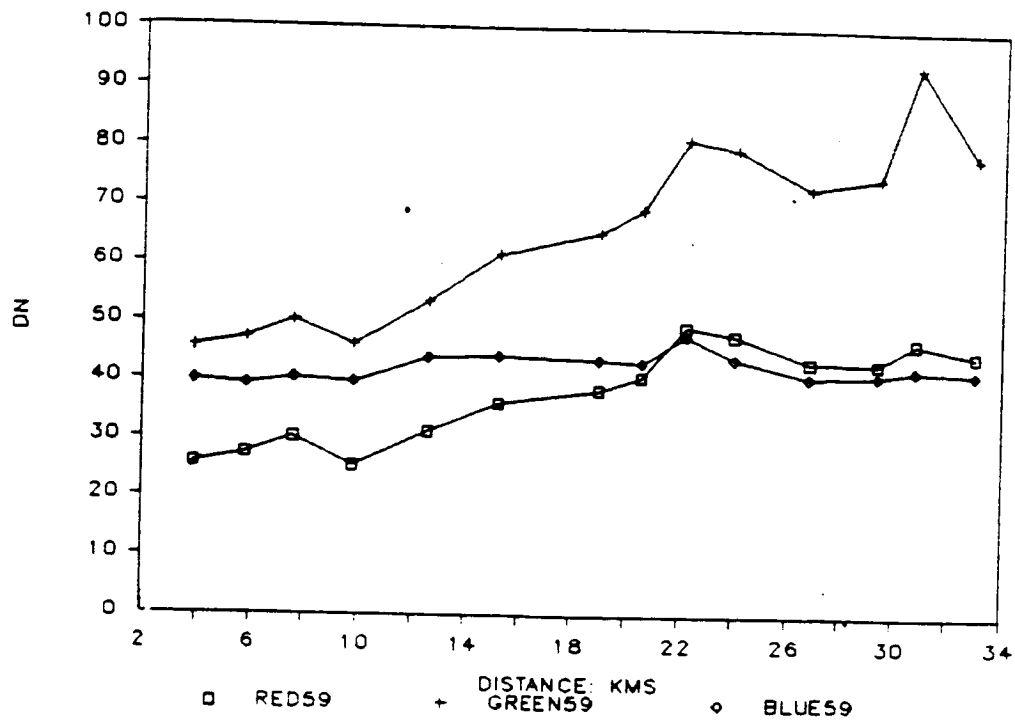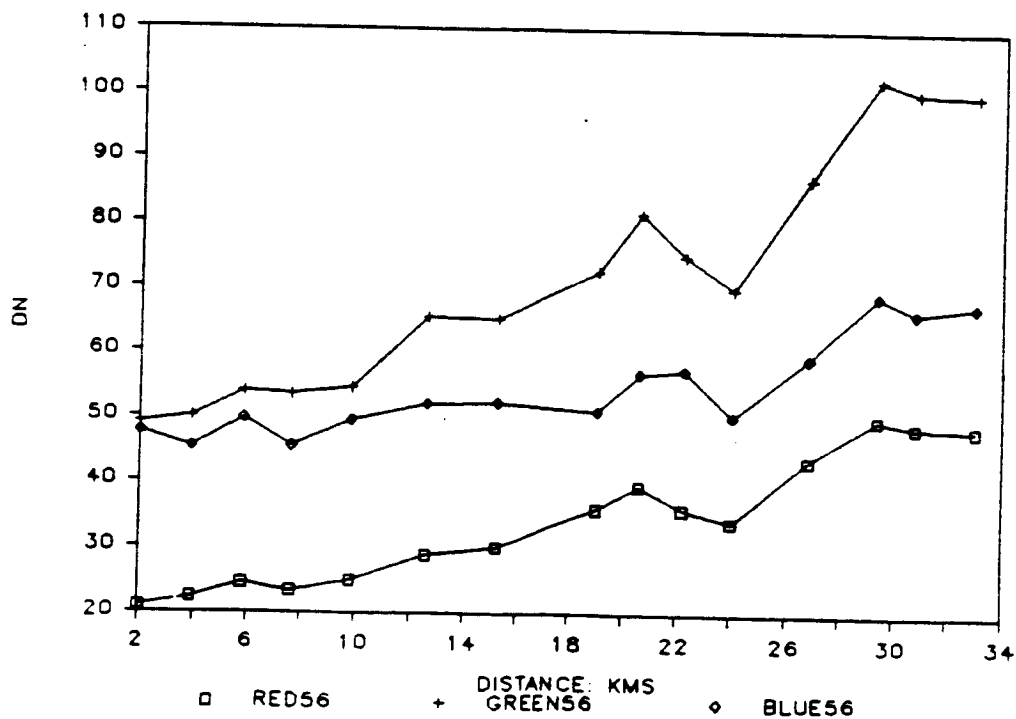
FRAME 59

Fig. 1

FRAME 56

Fig. 2

TABLE 2: SELECTED CORRELATIONS

1.

Frame	Distance	Elevation
Red 56	.97	.96
Green 56	.96	.96
Blue 56	.87	.91
Red 59	.90	-.82
Green 59	.94	-.88
Blue 59	.21	-.12

2.

Red 56	.97	-.96
Red 57	.98	-.95
Red 58	.98	-.95
Red 59	.90	-.82
Blue 56	.87	-.91
Blue 57	.98	-.97
Blue 58	.97	-.99
Blue 59	.21	-.12

3.

Red to Red:
    Frame 56-57:   .98
    Frame 56-58:   .97
    Frame 56-59:   .87

Blue to Blue:
    Frame 56-57:   .90
    Frame 56-58:   .89
    Frame 56-59:   .17

4.

Red 56 - Red 59:	.87
Green 56 - Green 59	.91
Blue 56 - Blue 59	.17

less substantive information. Obviously, this is an important question which subsequent research should address.

Figures 3-5 depict the color bands of each image based on distance. Considering red, Figure 3, frame 56 should show either the brightest or dimmest values (depending on the atmospheric effects) because of the extreme geometry. However, it's position beneath red and partially mixed with 57 cannot be explained. A general pattern of curves is evident, however (with 57 having some deviations), which is indicative of the flow's basic tonal configuration. There is good correlation between each frame's red and distance (subset #2 on Table 2) and good red-red relationships (#3), suggesting that a fairly consistent pattern of response in the red occurs under different viewing angle even when density values change. The same is true for blue except for frame 59, as evidenced above. Green reacts similar to red (not shown).

A second surprise results: the model of lava flow is to erupt in its most fluid state (lowest viscosity), cool as it travels downslope, and eventually becoming so viscous that it breaks into a jumbled jagged pile; i.e., from smooth, relatively reflective pahoehoe to broken, darker aa lava. Figures 3-5 show the opposite effect.
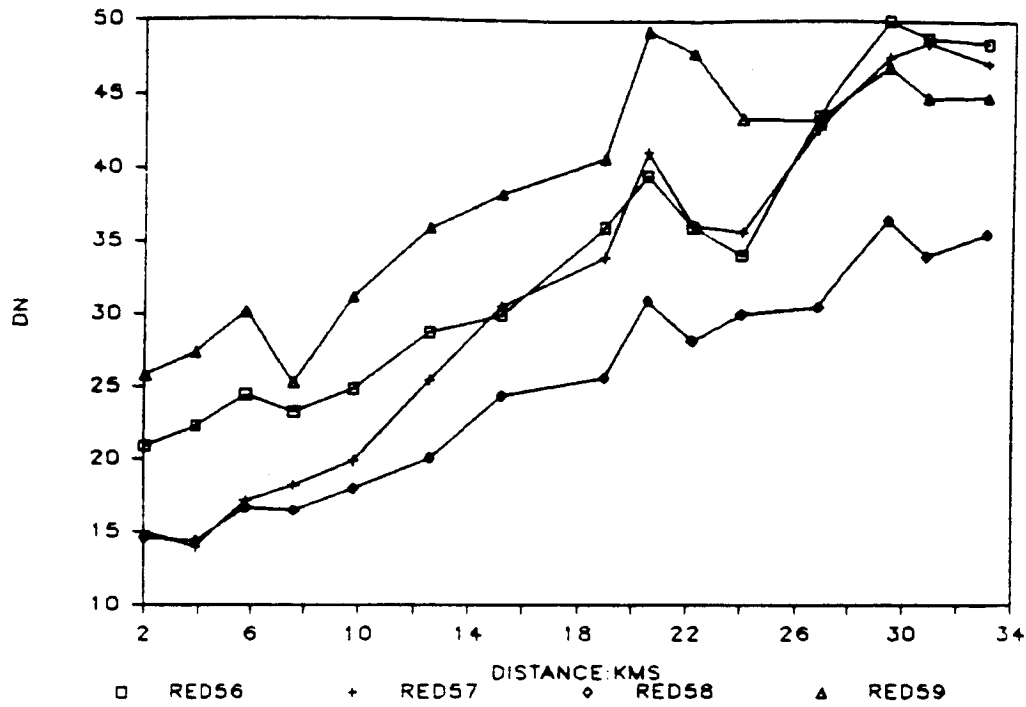
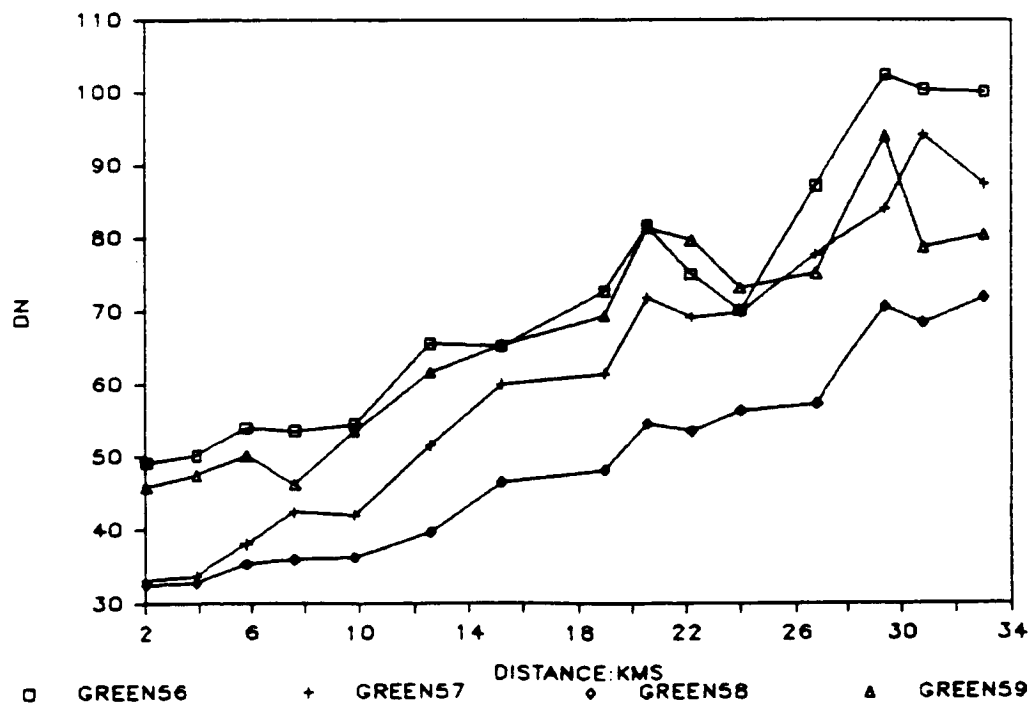# RED BANDS

Fig. 3

GREEN BANDS
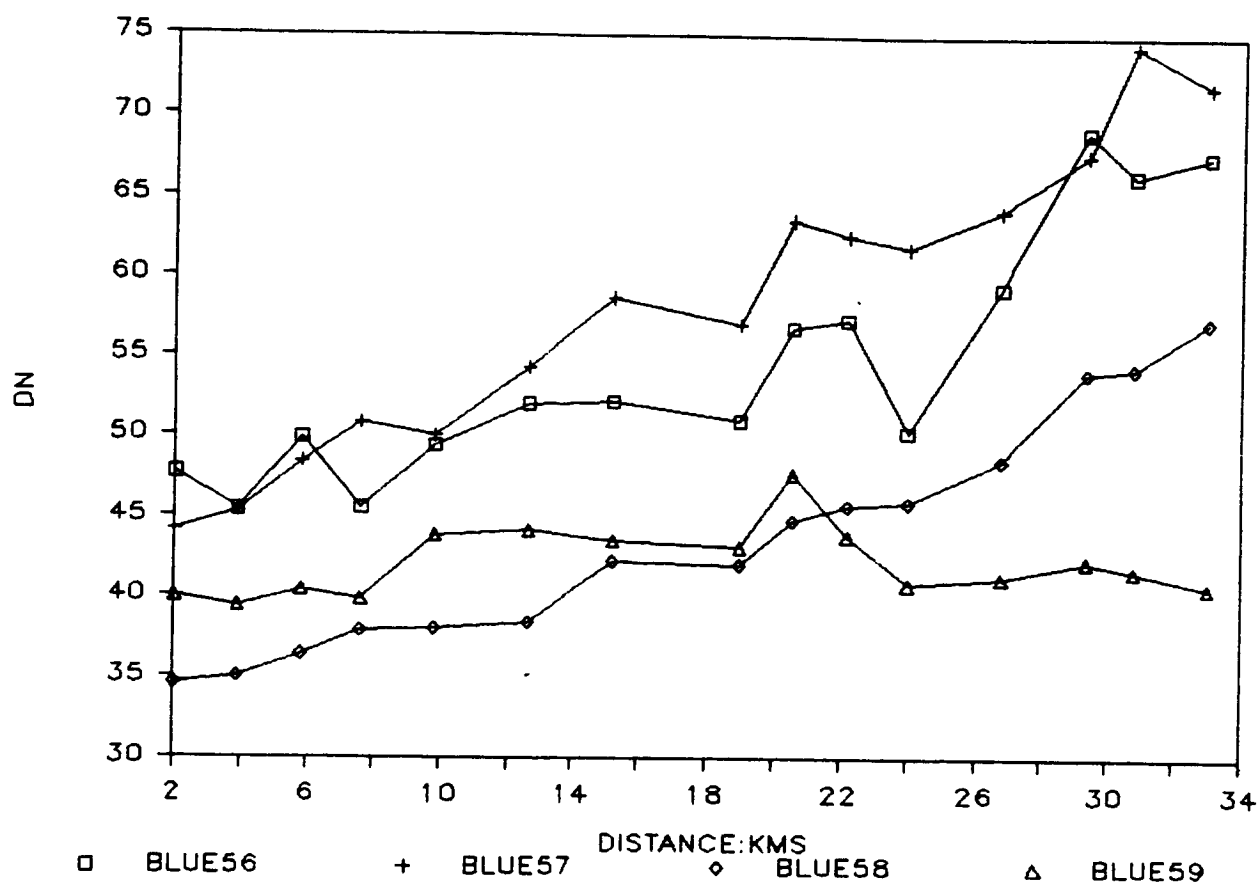

Fig. 4

BLUE BANDS

Fig. 5

Subset 4 in Table 2 presents a comparison between colors in the end frames. Red is a better penetrator of atmosphere than is blue, so a better correlation between is suggested. Analysis of residuals (deviations from regression lines) may offer further insight into the nature of the curves and comparisons of Figures 3-5.

Figures 6 and 7 display red and blue values as related to elevation. There is a -.98 correlation between distance and elevation, so the results are similar to those discussed above, except in an opposite direction. Specifically, darkest pixels are found in the upper elevations and get progressively brighter downslope. There is such a mix of lines in these graphs that consistency of viewing angle change responses is not apparent. However, note that there are similar tonal changes at particular elevations (e.g., a "bump" at 1600 meters), which can be useful for in-situ investigation, i.e., why is there a sudden change in the pattern of tones at a given elevation? Table 2 gives some band and elevation relationships, which, as expected, are very high except for frame 59's blue.
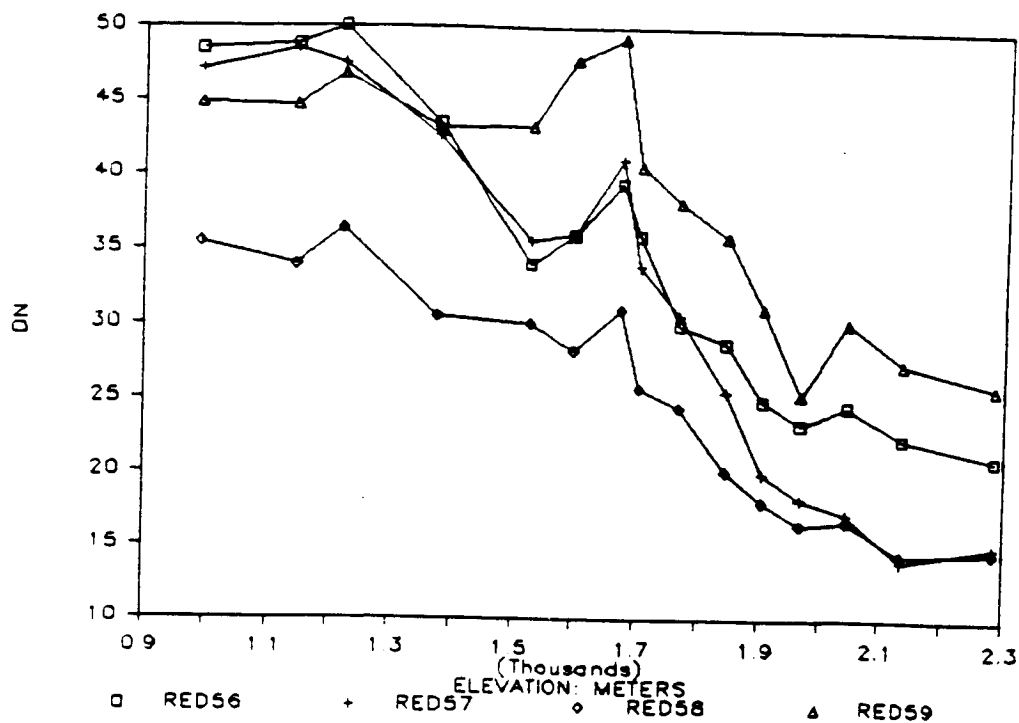
RED: ELEVATION

ELEVATION: METERS

□ REDS6   + REDS7   ◇ REDS8   ▲ REDS9

Fig. 6

BLUE: ELEVATION

ELEVATION: METERS

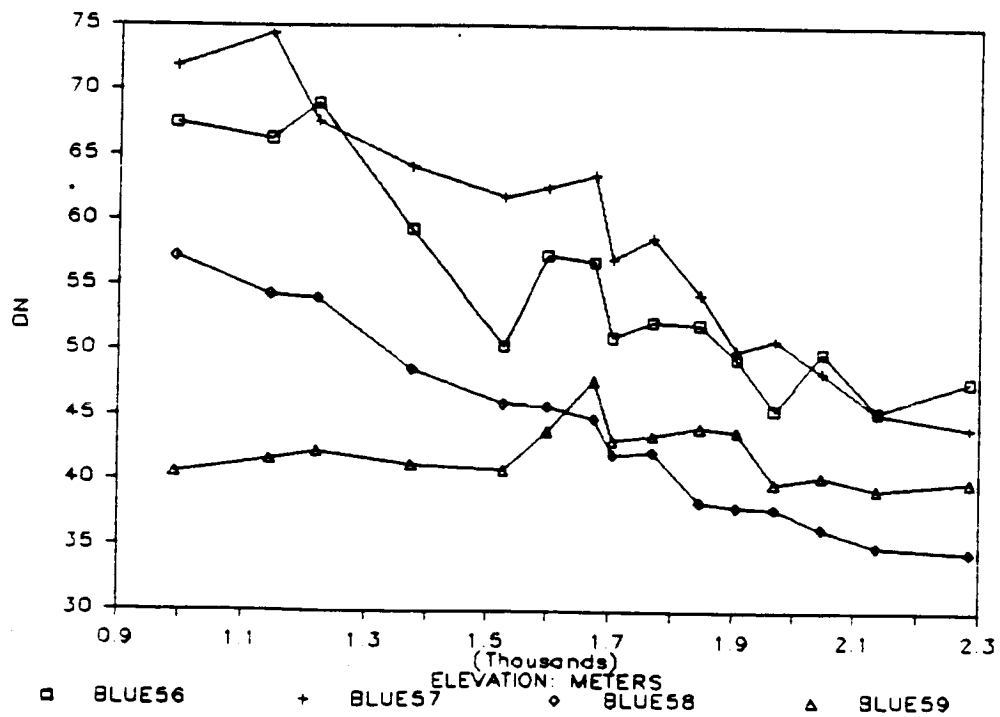□ BLUES6   + BLUES7   ◇ BLUES8   ▲ BLUES9

Fig. 7

CONCLUSIONS

Conclusions primarily address procedures; data analysis statements are initial thoughts rather than pure results. For brevity, the major results are listed, with comments.

Procedures:

1. The procedures developed in this project are, for the most part, measures needed to overcome problems that shouldn't have been present in the first place. However, they are simplified and manageable techniques--qualities often missing in digital image analysis reports and publications. While they are tedious and time-consuming, results were satisfactory and useful.

2. With improved software (see Problems and Suggestions number 3 below), some steps can be combined, shortened, or even excluded, making the basic outline more efficient and convenient. For example, direct application of Formula 2 would reduce work time by a great margin. Therefore, the methods developed here are preliminary and subject to further development.

3. With improved hardware (see Problems and Suggestions number 2), several of these steps may be eliminated altogether. Most of this project was spent dealing with equipment problems but with a matured, developed system, spectral anlysis of photographic imagery can proceed efficiently. There is very little active research on the use of digitized photograph under way, and apparently NONE for Shuttle photography. With a few improvements, JSC will have the capacity to undertake a useful, needed, and unique direction in remote sensing.

DATA

1. Although data are preliminary, there seems to be adequate information for productive analysis in the effects of off-nadir remote sensing of lava flows. Questions arise in attempting to explain why the theoretically-worst image presents information better than some others. Also, why is blue so poor from the supposedly best frame? Is there a quantitative trend of change from one viewing angle to the next? Nonetheless, data seem satisfactory and can be used to address these questions.

2. Tonal information offers insight into the morphologic nature of lava flows. Seen here is that the Keamuku flow

exhibits non-linear and unexpected tonal change downslope, indicating that the simplified model of flowing lava and attendant darkening is not applicable here. Many questions are raised by initial interpretation of data, such as reasons for corruption of the flow model, why specific locations deviate from the general pattern, etc. More analysis and much ground investigation are needed.

3. There are many statistical techniques that can be applied to imagery. Only a very few of the most simple have been used here, but the data are convenient and amenable to almost any technique desired.

Problems and Suggestions:

1. Access to equipment: Despite perfect relations with personnel and good intentions all around, access to the digitizer and computer were limited to the convenience of all other users. The short period that summer faculty fellows have for research and reporting make higher priority desirable and necessary. Perhaps an equable scheduling arrangement will evolve as the system is matured (which includes a move to Building 31). Given the extreme time lost in dealing with equipment problems, access time was even more valuable than under ordinary circumstances.

2. Equipment problems: Obviously, this project was greatly impacted by equipment failures and problems, most of which seem to be rectified, or at least manageable, at present. As the system matures, minor difficulties will be controlled and more personnel will be able to use it. The single critical suggestion that can be made is to replace the condenser lighting system with a diffusing type. Had one been in place for this project, progress would be far beyond that which was possible under the current problematic light source.

3. Procedures: While LIPS is an excellent image analysis system, it retains some limitations that created difficulties in this research. Currently, only a few people (at best) in the JSC-LPI area have internal access to LIPS and can write codes to make individual changes. Rebecca McAllister at LIP, for example, has written several new and enhanced routines, greatly improving LIPS' capabilities. The inability for most users to enter individual formula hinders production. Either a general routine for emplacing formulae or a system for having a knowledgeable programmer accomplish the job is needed.

BIBLIOGRAPHY

Bartlett, D., Johnson, R., Hardisky, M., Klemas, V. 1986. "Assessing Impacts of Off-Nadir Observation on Remote Sensing of Vegetation: Use of the Suits Model." International Journal of Remote Sensing 7,2:247-264.

Daughtry, C., Ranson, K. 1986. Measuring and Modeling Biophysical and Optical Properties of Diverse Vegetative Canopies. West Lafayette, IN: Laboratory for Applications of Remote Sensing, Purdue University. LARS Report 043086.

Duggin, M. J. 1985. "Factors Limiting the Discrimination and Quantification of Terrestrial Features Using Remotely Sensed Radiance." International Journal of Remote Sensing 6,1:3-27.

Egorov, V., Zhukov, B., Kottzov, V. 1978. "Analysis of Informative Characteristics of Scanner and Photo Images of The Earth Surface." Proceedings of an International Conference of Earth Observation from Space and Management of Planetary Resources. ESA Sp-134:455-464.

Gallegos, S., Nerem, R., Gray, T., Helfert, M. 1984. "Vegetative Responses From a Great Barrier Reef Surface Water Feature Detected by Space Shuttle Photography." Technical Papers: 1984 ASP-ACSM Fall Convention. Falls Church, VA: American Society of Photogrammetry/American Congress on Surveying and Mapping, 699-707.

Gerst S., Simmmer, C. 1986. "Radiation Physics and Modelling for ᵤrf-Nadir Satellite-Sensing of Non-Lambertian Surfaces." Remote Sensing of Environment, 20:1-29.

Goel, N., Deering, D. 1985. "Evaluation of a Canopy Reflectance Model for LAI Estimation Through Its Inversion." IEEE Transactions on Geoscience and Remote Sensing GE-23,5:674-684.

Holben, B., Kimes, D., Fraser, R. 1986. "Directional Reflectance Response in AVHRR Red and Near-IR Bands for Three Cover Types and Varying Atmospheric Conditions." Remote Sensing of Environment 19:213-236.

Hoffer, R., Anuta, P., Phillips, T. 1971. "ADP, Multiband and Multiemulsion Digitized Photos." Photogrammetric Engineering and Remote Sensing 38, 10:989-1001.

Jensen, J. 1986. Introductory Digital Image Processing: A Remote Sensing Perspective. Englewood Cliffs, NJ: Prentice-Hall.

Jensen, J., Estes, J., Tinney, L. "High-Altitude Versus Landsat Imagery for Digital Crop Identification." Photogrammetric Engineering and Remote Sensing 44,6:723-733.

Juday, R. 1986. "Spectrometry With Color Film (Exposed With No Spectral Isolation Filtering)." in STS 51-L JSC Visual Data Analysis Sub-Team Report: Volume II--Appendixes: D-131-139.

Kaltenbach, J., Helfert, M., Wells, G. 1984. "The View From the Shuttle Orbiter--Observing the Oceans From Manned Space Flights." Proceedings of SPIE--The International Society for Optical Engineering: Volume 489: Ocean Optics VII. Bellingham, WA: Society of Photo-Optical Instrumentation Engineers, 203-207.

Kimes, D., Newcomb, W., Nelson, R., Schutt, J. 1986. "Directional Reflectance Distributions of a Hardwood and Pine Forest Canopy." IEEE Transactions on Geoscience and Remote Sensing GE-24,2:281-293.

LeSchack, L. 1971. "ADP of Forest Imagery." Photogrammetric Engineering and Remote Sensing 37,8:885-893.

Li, X., Strahler, A. 1985. "Geometric-Optical Modeling of a Conifer Forest Canopy." IEEE Transactions on Geoscience and Remote Sensing GE-23,5:705-720.

Lord, D., Desjardins, R., Dube, P., Brach, E. 1985. "Variations of Crop Canopy Spectral Reflectance Measurements Under Changing Sky Conditions." [Photogrammetric Engineering and Remote Sensing 51,6:689-695.

McDowell, D., Specht, M. 1974. "Determination of Spectral Reflectance Using Aerial Photographs." Photogrammetric Engineering and Remote Sensing 40,5:559-568.

Nerem, R., Holz, R. 1984. "The Use of NOAA-n AVHRR Satellite Data and Hand-Held Earth Photography From Space in a Multi-Data Study of the Nile Delta." Technical Papers: 1984 ASP-ACSM Fall Convention. Falls Church, VA: American Society of Photogrammetry/American Congress on Surveying and Mapping, 722-736.

Norman, J., Welles, J., Walter, E. 1985. "Contrasts Among Bidirectional Reflectance of Leaves, Canopies, and Soils." IEEE Transactions on Geoscience and Remote Sensing GE-23,5:659-667.

Quirk, B., Scarpace, F. 1980. "A Method of Assessing Accuracy of A Digital Classification." Photogrammetric Engineering and Remote Sensing 46,11:1427-11431.

Scarpace, F. 1978. "Densitometry on Multi-Emulsion Imagery." Photogrammetric Engineering and Remote Sensing 44,10:1279-1292.

Scarpace, F., Quirk, B., Kiefer, R., Wynn, S. 1981. "Wetland Mapping From Digitized Aerial Photography." Photogrammetric Engineering and Remote Sensing 47,6:829-838.

Simmer, C., Gerstl, S. 1985. "Remote Sensing of Angular Characteristics of Canopy Reflectances." IEEE Transactions on Geoscience and Remote Sensing GE-23,5:648-658.

Slater, Philip. 1983. "Photographic Systems for Remote Sensing." in R. Colwell, ed., Manual of Remote Sensing: Second Edition, Falls Church, VA: American Society of Photogrammetry, Chapter 6:231-292.

Smedes, H, Linnerud, H., Woolaver, L., Hawks, S. 1971. "Digital Computer Mapping of Terrain by Clustering Techniques Using Color Film as a Three-Band Sensor." EG&G Technical Memorandum B-542, Bedford, MA: EG&G, Inc.

1986

NASA/ASEE Summer Faculty Research Fellowship Program

Johnson Space Center

University of Houston

Development of a Computer Program to Generate Typical
Measurement Values for Various Systems on a
Space Station

Prepared by: Louis A. DeAcetis, Ph.D.

Academic Rank: Professor of Physics

Institution: Bronx Community College of the
                    City University of New York
                    Bronx, NY 10453

NASA/JSC

Directorate: Engineering

Division: .   Tracking and Communications

Branch: Communications Performance and
                    Integration

JSC Colleague: Oron L. Schmidt

Date: August 15, 1986

Contract #:   NGT-44-005-803   (University of Houston)

Abstract

The elements of a simulation program written in Ada have been developed. The program will eventually serve as a data generator of typical readings from various Space Station equipment involved with Communications and Tracking, and will simulate various scenarios that may arise due to equipment malfunction or failure, power failures, etc.

In addition, an evaluation of the Ada language has been made from the viewpoint of a FORTRAN programmer learning Ada for the first time. Various strengths and difficulties associated with the learning and use of Ada are considered.

Introduction

The planning and testing of the various data col-
lecting and management systems on the proposed U. S.
Space Station requires that various configurations be
tested to determine whether performance criteria can be
realistically met. This requires that the various sys-
tems and their interaction with the data
collection/monitoring and data management systems be
simulated. (Eventually, as actual equipment becomes
available, it can replace its software simulation in
the monitoring system.) In particular, readings from
anticipated equipment involved with Tracking and Com-
munication are to be simulated by computer programs
which will mimic the typical performance of such equip-
ment, and, eventually, simulate the effects of anomalis
behavior or equipment failure.

It is anticipated that Ada will be used as the programming language for the simulation. Since there has been limited experience thus far with Ada, an evaluation of the language and its capabilities in this area is also of interest. The two goals of this project thus were:

a) To begin the programming of a data generator which will simulate various equipment involved with Tracking and Communications on the Space Station, and

b) To evaluate the Ada language as the language of choice for such a simulation.

Results and Conclusions

In order to meet the two major goals of this project, the principal investigator undertook to learn Ada, and then to apply it in programs that could serve as a basis for (or at least a start on) the required simulation. Headway was made in defining the parameters to be included in the simulation, and preliminary programs have been written that can serve as a basis for further programming.

In terms of an evaluation of Ada, the following comments about Ada are reported:

a) Ada is a language well suited for large scale simulation programs. In fact, it appears designed specifically for such applications (It would be cumbersome for small programs or numerical analysis).

b) Ada is complex and relatively difficult to learn (a background in PASCAL appears to be helpful). It contains much jargon, and many of the current references are poorly indexed.

c) The diagnostics on the VAX (typical?) are poor and not very helpful.

d) All of the bugs are apparently not worked out of the DEC implementation on the VAX-- we discovered one involving the TEXT_IO package and port I/O and were apparently the third facility to bring this to DEC's attention.

e) One of Ada's most powerful facilities is the
ability to program concurrent (or parallel) tasks that
execute "simultaneously." (Outside of Ada, such concur-
rent processing is only possible at the system level by
"detaching" from jobs that then continue to run in the
background. Such processes, however, are machine/system
dependent and therefore not readily transportable.)
This feature should be especially useful in programming
simulations.

The general conclusions of this project are that
Ada is a powerful language that is well suited to
programming simulations, but that its complexity means
that it is not for the casual programmer.

References

Booch, Grady, Software Engineering With Ada, Benjamin/Cummings Publishing Co., Menlo Park, CA, 1983.

Price, David, Introduction to Ada, Prentice Hall, Englewood Cliffs, NJ, 1984.

Reference Manual for the Ada Porgramming Language, Digital Equipment Corporation.

Telesoft-Ada Compiler User's Manual, Version 1.3G, Telesoft, 1983.

VAX Language-Sensitive Editor User's Guide, Digital Equipment Corporation, AA-FY24A-TE, July 1985.

VAX Ada Programmer's Run-Time Reference Manual, Digital Equipment Corporation, AA-EF88A-TE, February 1985.

Acknowledgements

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

JOHNSON SPACE CENTER

SPACE STATION ELECTRICAL POWER DISTRIBUTION ANALYSIS
USING A LOAD FLOW APPROACH

Prepared by:	Ervin M. Emanuel, P. E.
Academic Rank:	Assistant Professor
University and Department:	Prairie View A&M University Department of Electrical Engineering
NASA/JSC Directorate:	Engineering
Division:	Avionics Systems
Branch:	Control Subsystems
JSC Colleagues:	Bob Hendrix Kenneth J. Cox
Date:	August, 1986

SPACE STATION ELECTRICAL POWER DISTRIBUTION ANALYSIS
USING A LOAD FLOW APPROACH

Ervin M. Emanuel, P. E.
Assistant Professor of
Electrical Engineering
Prairie View A&M University
Prairie View, Texas

ABSTRACT

The Space Station's electrical power system will evolve and grow in a manner much similar to our present terrestrial electrical power system utilities. The initial baseline reference configuration will contain more than 50 nodes or busses, inverters, transformers, over-current protection devices, distribution lines, solar arrays, and/or solar dynamic power generating sources. The system is designed to manage and distribute 75 KW of power single-phase or three-phase at 20 KHz, and grow to a level of 300 KW steady state, and must be capable of operating at a peak of 450 KW for 5 to 10 minutes.

In order to plan far into the future and keep pace with load growth, a load flow power system analysis approach must be developed and utilized. This method is a well known energy assessment and management tool that is widely used throughout the Electrical Power Utility Industry.

This report will discuss and document the results of a comprehensive evaluation and assessment of an Electrical Distribution System Analysis Program (EDSA). Its potential use as an analysis and design tool for the 20 KHz Space Station electrical power system will be addressed.

1.0 INTRODUCTION

An orbiting Space Station employing a hybrid autonomous ac electrical power system will require an extensive computer aided analysis procedure for determining several operational characteristics of the system during transient and steady state conditions. A thorough analysis and smart control of earth based electric power systems have been achieved to a great degree by utilizing the "load flow" concept. Since the Space Station's electrical power system will evolve in a manner much like our own terrestrial electrical power system, the load flow approach should provide useful information in regard to the system present and future state of operation and load performance reliability.

2.0 LOAD FLOW CONCEPTS

Load flow is a computer-aided analysis procedure aimed at determining the actual power flow patterns (watt, vars) in a given system, and ways of controlling these patterns (1) J. Ward and Harry Hale of Iowa State University are are often credited with the first formulation of the power flow problem. Several other early and recent authors, such as G. Kron, R. L. Parks, G. W. Stagg, A. H. El-Abiad, J. R. Neuenswander, W. D. Stevenson, P. M. Anderson, K. C. Kruempel, and A. A. Fouad have made significant contributions directly or indirectly because of their interest in electric power systems.

2.1 OBJECTIVE OF LOAD FLOW ANALYSIS (LFA)

In order to determine the best ways of operating a given power

system and plan for future load growth, it is necessary to analyze the steady state solution of the network and achieve the following objectives.

● Determination of real and reactive power flow in the transmission line of a system based on certain prior assumptions regarding loads and generation.

● Computation of all system voltages.

● Identify all overloaded transmission lines. (Operating too close to the transmission line limit which could cause cable overheating).

● Rerouting of power in case of emergencies.

● Determine which load flow pattern will results in "optimum dispatch".

2.2 CONTROL OF REAL AND REACTIVE POWER FLOW

In order to successfully manage the transmission and distribution of electrical energy. The following objective must be met in a normal ac power system.

● Maintenance of real power balance (1)

● Control of frequency

● Maintenance of reactive power balance

● Control of voltage profile

● Maintaining an "optimum" generation schedule

● Maintaining an "optimum" power routing scheme.

2.3 "LFA" NOT A STANDARD CIRCUITS PROBLEM

Although load flow concept involves the use of many conventional

circuit analysis methods, it is not a standard circuits problem for several reasons. Some of these are:

- More often non-linear than linear due to the product of voltage and current being equal to power.

    - Additional non-linearities arise from the specification and use of complex voltage and current, and transmission components such as tap changing transformers in which the tap is adjusted to keep the bus voltage magnitude fixed.

    - Load impedance can never be represented by a constant due to a wide variety of load variations.

    - A typical load flow analysis involves network equations written in terms of voltages and powers (inherently non-linear and thereby demanding a numerical solution in most cases.)

    - A simple two bus one-line load flow diagram is shown in Figure 1 depicting the variable of interest in a load flow analysis.

3.0 SPACE STATION LFA

Space Station load flow analysis in theory will demand the same principle requirements as earth-based system. However, the practical control aspects of this second generation multisource hybrid power system will require considerable attention. Some of these include generation and frequency control of large area solar arrays and solar dynamic generating sources.

Other areas of concern are:

- Accurate modeling of all system components (i.e., transmission and distribution lines, transformers, etc.)

- Resonance conditions observed when using litz cable for a 20 KHz power system (2).

- Variable network topology due to a random distributed load mixture.

- Hybrid load flow numerical methods development.

The reference configuration electrical distribution system is shown in Figures 2 and 3.

4.0 ELECTRICAL DISTRIBUTION SYSTEM ANALYSIS (EDSA)

EDSA is an integrated collection of computer programs for electrical distribution systems analysis and design. The program is written to assist the engineer in the design and analysis of electrical power transmission and distribution systems of public utilities and large installations using 60 Hz power.

The programs are menu driven, interactive, and requires an IBM PC computer, 640 KB of memory capacity, and a 2.1 DOS or greater disk operating system.

Calculations are based on NEMA, ANSI, NEC, IEEE, Beaman, Stevenson, GE, and Westinghouse T&D reference books.

The complete EDSA software package includes the following programs:

- Short circuit analysis loop/radial/utility/multi-generating source. 185/300 bus system

- Load flow analysis 185/300 bus system

- Motor starting voltage dip load flow method

- Ground mat analysis

- Protective device coordination

- Shielding effectiveness

- Motor torque simulation

- Generator set sizing

- Wire and conduit sizing

- Transformer sizing

- Symmetrical ampere to withstand rating conversion factors

- Motor starting voltage dip impedance division method (on generator/utility)

- Panel, MCC, primary SWG, unit substation, automatic transfer switch, bus duct schedules

- Capacitor reactor starting of large synchronous motors

- Calculating zero-sequence resistance and reactance $R(0)$, $X(0)$

- Load flow multi-source and loop

- Computer graphics interface

- Build, edit feeder and transformer master data files

- Build, edit fuse relay and MCP data files

- Build circuit breaker data files

The data generated by the load flow and short circuit analysis programs are shown in Figures 4 and 5.

5.0 CONCLUSIONS

The Space Station electrical power system will be a complex, high power, multisource system operating with a high degree of autonomy. The simulated management and control of this second generation system will require a number of artificial intelligence and expert system

software modules. Some of these are: load flow analysis, state
estimation, generation control, and energy storage.

The load flow analysis component of the EDSA package proved to be
very useful for analyzing a 20 KHz electrical distribution system.
This software package is considered by many to be the most complete
analysis and design tool available.

References

1. Elgerd, O. I.: <u>Basic Electric Power Engineering</u>. Addison Wesley Publishing Company, 1977.

2. Raley, J. B., and Piechowski, M. T.: "High Frequency Power Transmission Line Characteristic And Applications In Space Station Power Distribution Systems," Space Station Definition and Preliminary Design, (WP-02), NAS9-17367, April 1986.

3. Heydt, G. T.: <u>Computer Analysis Methods For Power Systems</u>. MacMillan Publishing Company, 1986.

4. Stagg, G. W., and El-Abiad, A. H.: <u>Computer Methods In Power System Analysis</u>. McGraw-Hill, Inc., 1968.

5. Neuenswander, J. R.: <u>Modern Power Systems</u>. International Text Book Company, 1971.

6. Anderson, P. M.: <u>Analysis Of Faulted Power Systems</u>. The Iowa State University Press, 1973.

7. Lazar, Irwin: <u>Electrical Systems Analysis And Design For Industrial Plants</u>. McGraw-Hill Book Company, 1980.

8. NASA Conference: "Space Power Subsystem Automation Technology," Marshall Space Flight Center, Huntsville, Alabama, October 1981.

9. Emanuel, Ervin M.: "Digital Readout Of Polyphase Induction Motor Test Values," M. S. Thesis, Iowa State University, 1976.

10. Emanuel, Ervin M.: "Apparatus For Testing The Performance Of Electrical Machines," U.S. Patent No. 4091662 (1978).

11. Emanuel, Ervin M.: "A Digital Instrumentation System For Test And Control Of AC Machines," Proceedings of the 1978 Midwest Power Symposium, University of Nebraska, Lincoln, Nebraska, 1978.

12. Emanuel, Ervin M.: "Apparatus And Method For Testing The Performance Of Electrical Machines," U.S. Patent No. 4348892 (1982).

13. Emanuel, Ervin M.: "An IBM PC Based Math Model For Space Station Solar Array Simulation," NASA/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program, NCR 171931, August 1985.
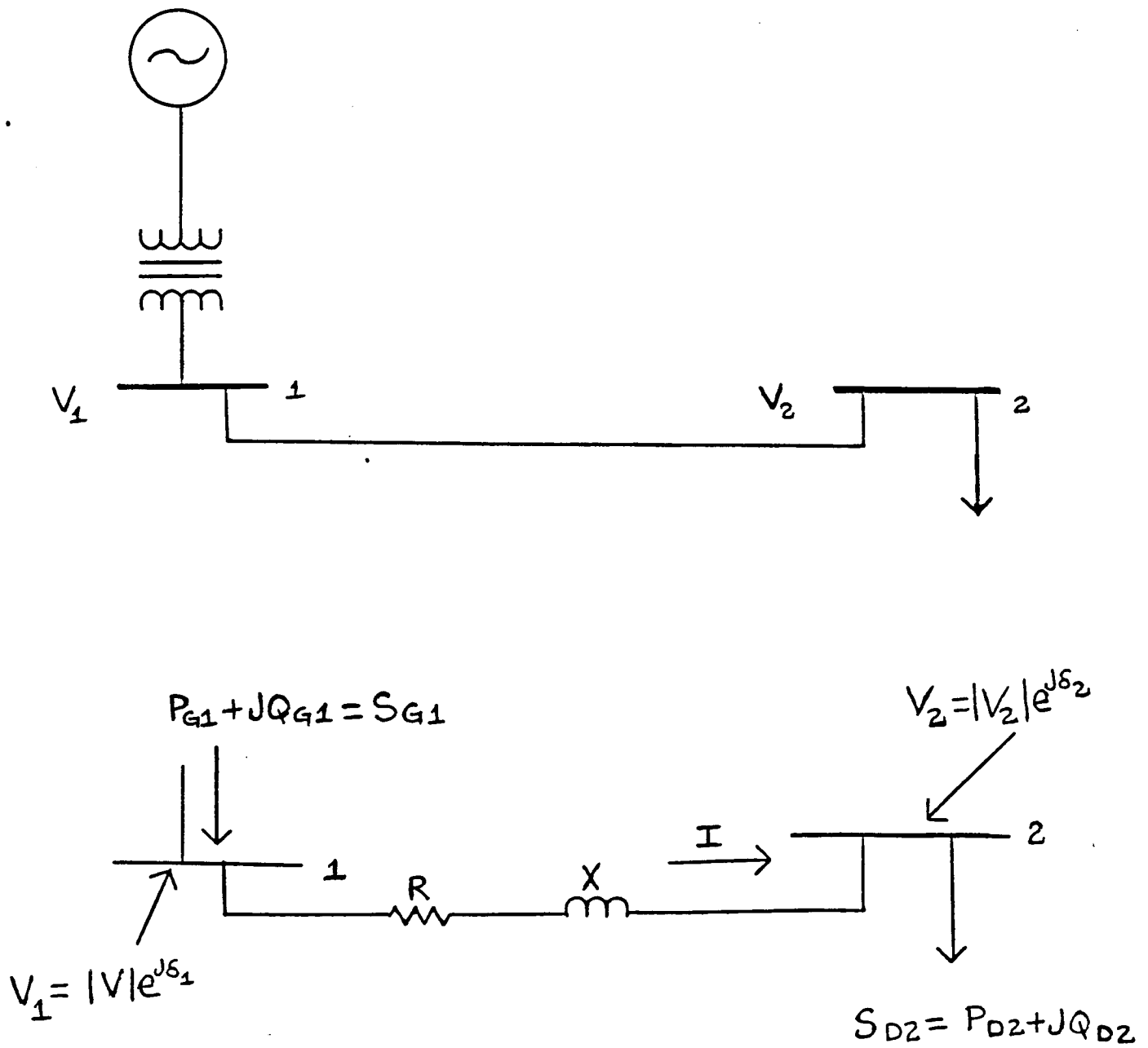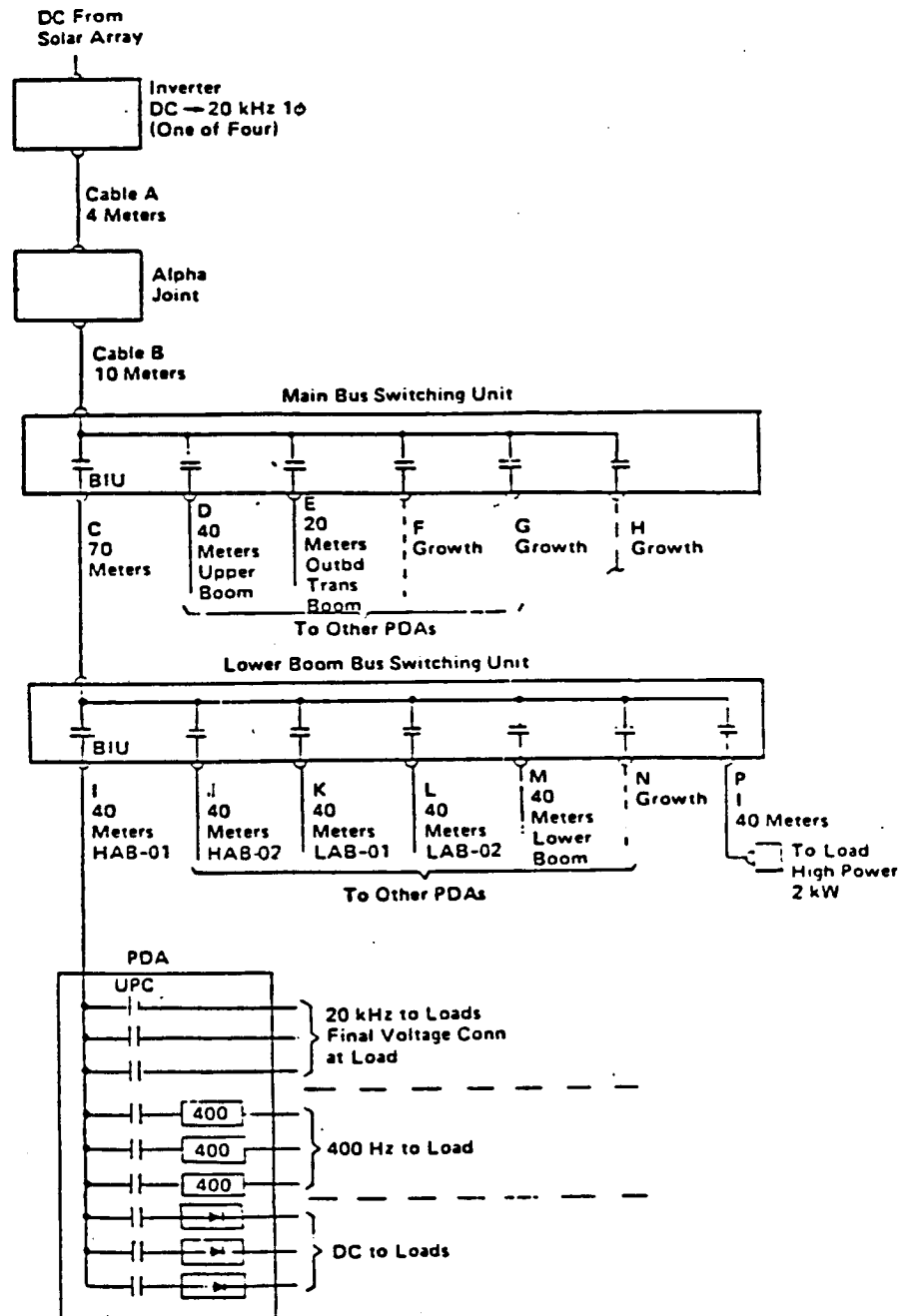
V_1 1 V_2 2

$P_{G1} + JQ_{G1} = S_{G1}$

$V_2 = |V_2| e^{J\delta_2}$

1   R   X   I   2

$V_1 = |V| e^{J\delta_1}$

$S_{D2} = P_{D2} + JQ_{D2}$

FIGURE 1.   ONE LINE LOAD FLOW DIAGRAM

FIGURE 2. REFERENCE CONFIGURATION ELECTRICAL DISTRIBUTION SYSTEM

DC From
Solar Array

Inverter

Cable A
4 meters

Alpha
Joint

Cable B
10 meters

Main Bus Switching Unit

Cable D
40 meters

Cable E
20 meters

Cable C
70 meters

Open
Circuit

Open
Circuit

Lower Boom Bus Switching Unit

Cable I
40 meters

Cable J
40 meters

Cable K
40 meters

Cable L
40 meters

Cable M
40 meters

Open
Circuit

Open
Circuit

Open
Circuit

Open
Circuit

These Two Cables
Will Be Loaded
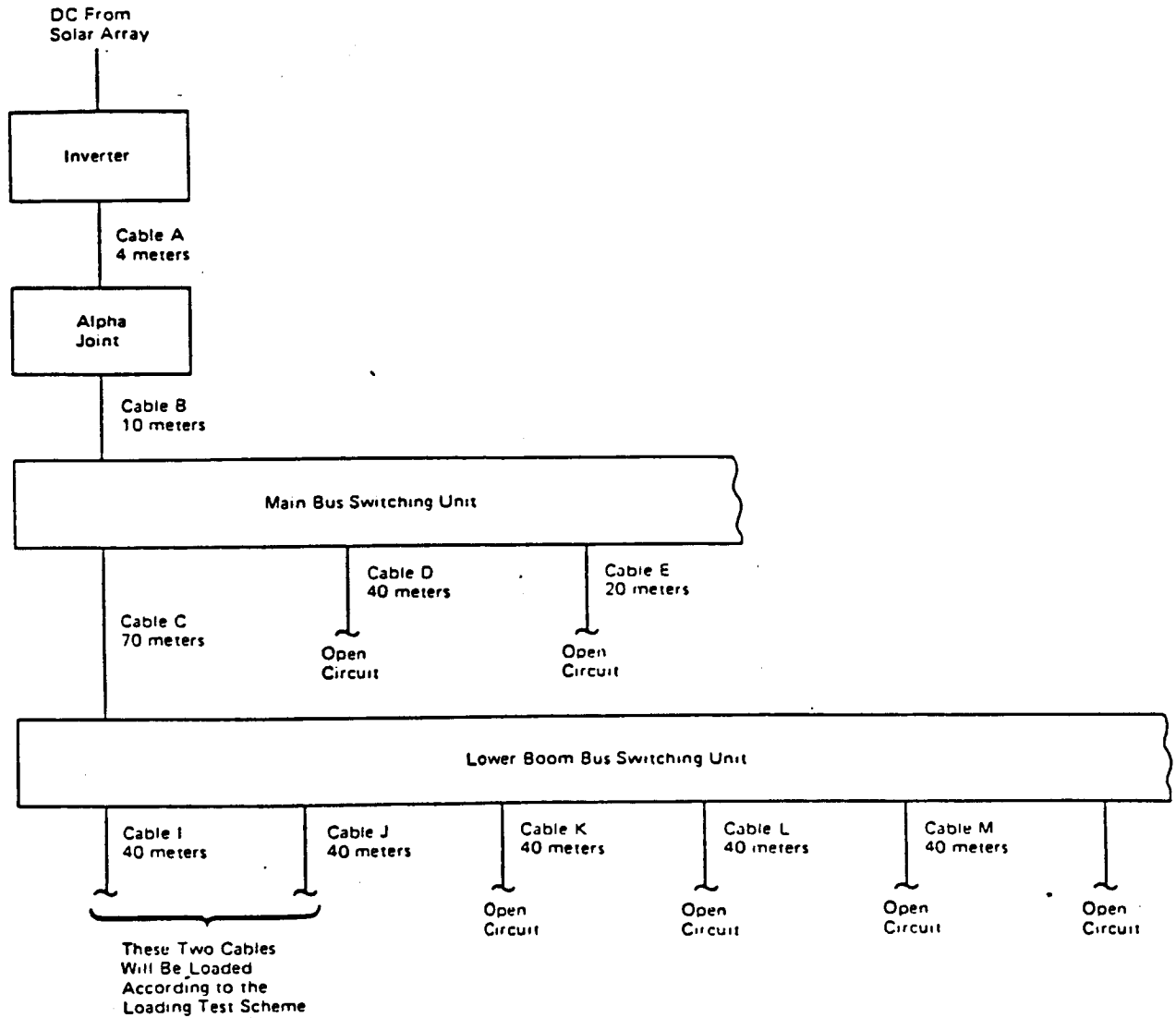According to the
Loading Test Scheme

FIGURE 3.   DISTRIBUTION SYSTEM BLOCK DIAGRAM

7-13

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

Texas A&M University

Active Vibration Control in Microgravity Environment

Prepared by: Carl H. Gerhold, PHD, PE

Academic Rank: Associate Professor

University & Department: Texas A&M University

                                 Department of Mechanical Engineering

NASA/JSC

Directorate: Engineering

Division   Structures and Mechanics

Branch   Loads and Structural Dynamics

JSC Colleague: A. Rodney Rocha

Date   August 8, 1986

Contract #:   NGT-44-005-803

Active Vibration Control in Micro-gravity Environment

Carl H. Gerhold, PhD, PE

Associate Professor
Mechanical Engineering Department

Texas A&M University
College Station, Texas 77843

The low gravity environment of the Space Station is suitable for experiments or manufacturing processes which require near zero-g. Such experiments are packaged to fit into rack-mounted modules approximately 106.7 cm (42 in) wide x 190.5 cm (75 in) high x 76.2 cm (30 in) deep. The mean gravitation level of the Space Station is expected to be on the order of 10^{-6} g (9.81×10^{-6} m/s^2). Excitations, such as crew activity or rotating unbalance of nearby equipment can cause momentary disturbances to the vibration-sensitive payload on the order of 0.4 g. Such disturbances can reduce the micro-gravity environment and compromise the validity of the experiment or process. Isolation of the vibration-sensitive payload from structure-borne excitation is achieved by allowing the payload to float freely within an enclosed space. Displacement-sensitive transducers indicate relative drift between the payload and the surrounding structure. Small air jets provide a negative thrust vector which keeps the payload centered within the space. The mass flow rate of the air jets is controlled such that the resultant acceleration of the payload is less than a criterion level of 10^{-5}g. It is expected that any power or fluid lines that connect the experiment to the Space Station structure can be designed such that they transmit vibration levels within the criterion. A flexible coiled hose such as is used to carry shop air has the requisite compliance.

An experiment has been fabricated to test the validity of the active control process and to verify the flow and control parameters identified in a theoretical model. Zero-g is approximated in the horizontal plane using a low-friction air-bearing table. An analog control system has been designed to activate calibrated air-jets when displacement of the test mass is sensed. The experiment demonstrates that the air jet control system introduces an effective damping factor to control oscillatory response. The amount of damping as well as the flow parameters, such as pressure drop across the valve and flow rate of air, are verified by the analytical model.

NASA Colleague: A. R. Rocha/ES4/x-4393

Active Vibration Control in
Microgravity Environment

Introduction

Microgravity research experiments, such as crystalline growth in zero
gravity, require a static gravitational environment on the order of 10^{-6} g.
Such a low gravity environment is obtained in space near the centerline of
a Space Station. In a low gravity environment, momentary disturbances,
such as thruster fire or crew pushoff introduce shocks to the vibration
sensitive experiment on the order of 3.0 x 10^{-3} g. [1] Rotating equipment,
such as pumps located near the experiment, may transmit steady state
vibration of 0.4 g at 10 Hertz. [2] Such disturbances alter the
microgravity environment and can degrade the validity of the experiment.

The purpose of this project is to devise and evaluate a method to reduce
the vibration transmitted from the Space Station structure to the vibration
sensitive payload. The most effective way to decouple the payload is to
allow it to float freely in space. Figure 1 is a schematic representation
of the payload suspended within an enclosure which is attached to the Space
Station structure. Unbalanced forces, whether generated internally or
externally, cause the payload to drift toward the enclosure. This relative
motion is sensed and used to activate air jets. The jets provide thrust to
stop the motion and to return the payload to a central location within the
enclosure.

External forces arise from gravity gradient over an orbit cycle and
aerodynamic drag-induced deceleration of the Space Station, and result in
long period drift between the enclosure and the payload. External forces
are also transmitted through connections such as air, fluid, or power
lines. The lines act as compliant elements which transfer structural

vibration to the payload. Unbalanced forces are generated internally by rotating equipment or fluid motion within the payload.

The air jet control system is required to keep the payload from contacting the enclosure. The criterion for the air jet is that the thrust produced by the jet results in a net acceleration of the payload less than or equal to 10^{-5} g.

The sensors used to control the jets respond to velocity and displacement of the payload. This type of control is expected to be sufficient for the following reasons. It is assumed that the internally generated forces are sufficiently small that the acceleration produced by them is less than 10^{-5} g. However, these forces may cause the payload to drift, which drift the air jet is intended to control. Gravity gradient is expected to be the major long period external excitation, producing a disturbance on the order to 10^{-5} g [1]. The period of this excitation is approximately 90 minutes (one orbit cycle). Since the gravity gradient does not exceed the allowable acceleration criterion, the air jet control which limits long term drift of the payload keeps it essentially neutrally buoyant. Structural vibrations of magnitude 0.4 g at 10 Hertz may be transmitted through hoses and power lines to the payload. Compliant connections, such as self-coiling flexible air lines, can be used to reduce the transmitted vibration so that the steady state acceleration of the payload is on the order of 10^{-5} g. The air jet control is used to limit drift resulting from this disturbance.

Preliminary investigations, including computer simulation, indicate that the air jet control system is feasible. However, in order to establish a workable system it is necessary to test the concept experimentally. The experimental setup consists of a mass constrained to move in a one-dimensional simulated low-g environment. The displacement and velocity of the mass are monitored and used to control solenoid·activated air valves. The first phase of the experimental program is intended to

establish the feasibility of the air jet feed back control system concept and to identify parameters for the air jet and the feedback control systems.

Theoretical Background

The basic one-dimensional model is derived from the system shown in figure 1. The one-dimensional model is felt to provide sufficient detail to identify system parameters. It is assumed that the internal forces generated by rotating unbalance or fluid motion within the payload are negligible in comparison to the force transmitted through the compliant coupling to the structure. This couple between the payload and the structure is modeled as a massless spring element. The differential equation describing the motion of the payload is:

$$M\ddot{x} = k(y-x) + F_j \tag{1}$$

where:

M = payload mass
k = equivalent stiffness of the hose or power line
x = absolute displacement of the payload
$\ddot{x}$ = absolute acceleration of the payload
y = displacement of the structure
F_j = thrust exerted by the jet

1. Jet Thrust

   The thrust exerted by the jet is modeled from basic fluid dynamics theory [3] as:

$$F_j = \frac{dm}{dt} (\dot{x} - v_f) \tag{2}$$

where:

$\dfrac{dm}{dt}$ = mass flow rate of the air jet

$\dot{x}$ = velocity of the payload

v_f = local flow velocity of the jet

The thrust produced always opposes motion of the mass. When the mass displacement is positive and to the right, the right side jet is activated to produce a left pointing thrust vector. Similarly, when the payload is to the left of the central position and moving toward the left, the left side jet is activated to produce a right pointing thrust vector. This model of the jet thrust has been verified in a static test performed at the Shock and Vibration Laboratory at Texas A&M University. The experiment consisted of a 1.41 kg mass suspended by wire 0.686 m. long. The air jet impinges on the mass and the angle the wire makes with the vertical is measured, as shown in figure 2a and 2b. Summing forces in the vertical and horizontal direction, the expected force balance is:

$$Mg \; \frac{\sin \theta}{\cos \theta} = \frac{dm}{dt} \, V_f \qquad\qquad (3)$$

where:

M = mass of pendulum

θ = static angle

$\dfrac{dm}{dt} \, V_f$ = momentum flux of the jet

The jet diameter is 9/32-in. (7.1 mm). The air supplied was compressed air used throughout the building. The flow rate was varied from 2.12×10^{-3} m^3/S (4.5 cfm) $7.55 \times 10^{-3} m^3/S$ (16.0 cfm). The results of the experiment

are shown in figure 3 in which the angle reached by the pendulum is plotted against flow. The relatively simple theory provides reasonable estimate of flow rate required to produce a thrust force on the mass. The figure shows that as the distance between the jet and the mass increases (angle increases), the flow rate required to maintain equilibrium is greater than the simple theory estimates. Factors contributing to this are experimental error; the fact that the air jet impinges on the mass at an angle of inclination, as shown in figure 2b, the thus the momentum flux is transferred less efficiently as the angle increases; and the loss of momentum flux due to temperature changes as the distance downstream of the jet exit increases. While experimental error and loss of momentum flux transfer due to impingement angle are factors particular to the static experiment, the possible loss of momentum flux due·to heat transfer with the surrounding air can affect the thrust in the active control project as is indicated in the following section.

2. Jet Flow Equations

The expressions for $\frac{dm}{dt}$ and V_f are derived based on the assumption that the momentum flux is constant throughout the flow field. It is required to know the mass flow rate and flow velocity separately because the thrust term in equation 2 depends on the relative velocity between the jet and the mass. The derivations for the terms are shown in Appendix A. For a one-dimensional, isothermal jet, the mass flow rate is:

$$\frac{dm}{dt} = 0.234 \times (M_o \rho)^{1/2} \tag{4}$$

where:

x = distance downstream of jet exit
M_o = momentum flux at the jet exit
ρ = density of air

The momentum flux is assumed constant and equal to the momentum flux at the jet exit, M_0, where, for a round jet:

$$M_0 = \rho \pi U_0^2 R^2 \tag{5}$$

where

U_0 = air velocity at the jet exit (assumed uniform)
R = jet radius

The mass flow rate is evaluated at a downstream location using equation 4, and since the momentum flux is constant, the equivalent uniform flow speed is calculated.

3. Expected Air Consumption Parameters

The thrust required to give a 45 kg mass an acceleration of 10^{-5} g is estimated to be generated by a flow rate of 2.24 x 10^{-5} m^3/S (0.048 cfm) through a jet of diameter 0.79 mm (1/32 in.). The pressure drop across the valve required to produce this flow is calculated [4] to be 3.45 x 10^3 N/m^2 (0.5 psi). An example of a compressor that can supply such a system is a 12 V DC, 187W (1/4 HP) oil free, piston compressor rated a 3.30 x 10^{-4} m^3/S (0.7 cfm) at 6.90 x 10^5 N/m^2 (100 psi). This is not expected to be a prohibitive power requirement, even if the control system operates continuously.

The type or frequency of excitation in the Space Station is not known. However, in order to assess the expected performance of the air jet system, the following example is investigated, based on the differential equation of motion:

$$M\ddot{x} = k\,(y-x) + \frac{dm}{dt}\,(\dot{x} - V_f)$$

where the terms are defined in the previous section.

The payload mass is 45 kg and the jet uses air at $21^{\circ}C$ ($70^{\circ}F$) and 1.0×10^{5} N/M^{2} (14.7 psi), with a flow rate of 2.24×10^{-5} m^{3}/S through a jet of 0.79 mm diameter. The spring stiffness is 5.0 N/m (0.34 lb/ft). This is the measured stiffness of a 95.3 mm (3.75 in) coil diameter spiral, self-coiling, flexible hose. The structural displacement, y, is assumed to be sinusoidal with period of 0.1/sec. The magnitude is calculated from the acceleration magnitude of 0.4 g. The jets are spaced such that the excursion of the mass is limited to $\pm$ 25.4 mm ($\pm$ 1.0 in). The control mechanism thresholds are set such that the jet is activated when the mass displacement exceeds 2.54 mm (0.10 in) or the velocity exceeds 1.0×10^{-4} m/S (3.94×10^{-3} in/S).

The mass is displaced 10.0 mm (0.40 in) and released from rest at t = 0. Without the air-jet control systems, the mass oscillates with small amplitude at the driving frequency superimposed on a vibration at the natural frequency of the spring-mass system. The peak acceleration of the payload due to transmissions of the 0.4g structural acceleration through the spring is calculated to be 1.13×10^{-5} g. The natural frequency vibration component has a period of 18 seconds, based on the mass and stiffness values, and amplitude equal to the initial displacement. In the absence of any control, this oscillation continues indefinitely.

With the air jet control, the response is shown in figure 4. The resultant motion of the mass is a sinusoid that decays linearly in time. The period of oscillation is 18 seconds, which corresponds to the natural period based on the spring and mass. The decay rate corresponds to an equivalent viscous damping factor of 0.0282. The time required for the mass to reach a steady state vibration about the central location is approximately 120 seconds. Of that time, the jet is on for 48.3 seconds, or 40 percent.

Experimental Program

The first phase of the experiment is intended to demonstrate the feasibility of the air jet control and to establish the jet flow parameters.

The experimental setup consists of the following major elements: a. sensor, b. electronic control system, c. air jets, and d. test mass. The electronic control system and air jets are designed for Space Station application. The sensor is a commercially available Linear Variable Differential Transforms (LVDT). This transducer limits allowable displacement to one dimension and thus is not applicable to Space Station application where three degrees of translation and three degrees of rotation are possible. The test mass is supported by an air-bearing table and is constrained to move in the horizontal plane. The air-bearing facility simulates zero gravity in the horizontal plane. The experimental set up is shown schematically in figure 5.

1. Control Algorithm

The LVDT produces a voltage which is proportional in magnitude and in sign to the displacement of the test mass. The voltage output from the LVDT is differentiated, producing a voltage signal proportional to the velocity of the test mass. The displacement and velocity proportional signals are each compared to thresholds values. The purpose for the threshold is to permit a dead band in which no control is activated. If the threshold is exceeded, a +15 volt signal is output on the line corresponding the sign of the input voltage, and a -15 volt signal is output on the other line. The outputs of the threshold detectors are combined in the comparator circuit. If the combined voltage on one of the lines is large (30v) and positive, this indicates that the mass is displaced from the center and moving further away. The comparator opens the relay to activate the appropriate jet. At the same time, the timer circuit is activated which limits the duration of the air jet pulse. If the combined voltage at the comparator is large and negative, this indicates either that the mass is within the

dead band or that the mass is displaced from the central position but is tending toward it. In either of these cases, no air thrust is required.

2. Experiment Parameters

The air-bearing table is intended to provide friction-free horizontal motion. Any friction at the air bearing surface will add damping to the system which degrades the validity of the air jet efficiency determination. In an effort to quantify the air bearing equivalent damping, an experiment was run using an air bearing pad on a laboratory quality marble slab at the Shock and Vibration Laboratory at Texas A&M. The pad was connected to ground by springs, loaded with 90.72 kg (200 lb) and set into free vibration. From the logarithmic decrement, the damping coefficient, c, was measured to be 0.342 Ns/m (1.95×10^{-3} lb-s/in). The experiment was repeated vertical plane to eliminate the air friction. The damping coefficient was again found to be 0.342 N-s/m. Thus, the air film damping is negligible in comparison to the internal damping of the springs. The expected valve of the viscous shear damping coefficient is calculated from:

$$c_{exp} = A\mu/h$$

where:

A = contact area of the bearing surface
μ = dynamic viscosity
h = film thickness

The bearing film thickness is on the order of 0.051 mm (0.002 in). The expected damping coefficient is c_{exp} - 8.38×10^{-3} N-s/m (4.79×10^{-5} lb-s/in) for a bearing 0.152 m x 0.152 m (6 in x 6 in). The theoretical coefficient is 2.5 percent of the measured coefficient, which includes the springs and air bearing together. The theoretical value does not account for turbulent flow or surface roughness. Thus, as a first approximation,

it is assumed that the air bearing damping is 10 percent of the measured coefficient, or 3.42×10^{-2} N-s/m. The system defined in a previous section consists of a 45 kg mass connected to a spring with stiffness 5.0 N/m. The air jet control introduces an equivalent damping factor of 0.0282. The expected damping factor for three air bearing is 1.14×10^{-3}. Thus, the damping introduced by the air bearing is expected to be negligible in comparison to the damping introduced, based on the damping coefficient assumed above by the control system.

The air-jet momentum flux, M_o, required to produce an acceleration of 10^{-5} g of a 45 kg mass is 4.41×10^{-3} N. The jets used in this experiment are commercially available solenoid operated air valves fitted with plugs in which a 0.799 mm (0.03125 in) hole has been drilled. The force exerted by the flow from the jet was measured statically by impinging the flow on a scale. The force versus pressure ratio across the jet curve is shown in figure 6. It is found that the force decreases linearly as the distance from the jet to the scale decreases. This indicates that the assumption of constant momentum flux is incorrect. However, the overall percent difference from the lowest to the highest force is approximately 25 percent. Thus, the constant momentum flux assumption is a valid first approximation. The curve of expected force is shown in the figure. The expected curve is derived from sharp-edged orifice flow theory [4], and is found to provide a good estimate of the force. The transient response of the jet was measured using a hot-wire anemometer. When the switch activating the solenoid is closed, the response shows second order characteristics with approximately 1.5 percent overshoot reached at 220 m-sec. This response time is a combination of the air-jet response and the anemometer response. The anemometer response time was measured separately and found to be approximately 150 m-sec. Thus, the response of the anemometer is dominant in the total measured system response. As a first approximation, the response time of the valve is assumed to be the difference, or 70 m-sec. The air-jet, when activated, will pulse for a predetermined time period of 0.50 sec. Thus, the response time is expected

to have a negligible effect on the active control system performance.

Results

The one-dimensional test set up has been fabricated and assembled on the precision air-bearing floor in the Technical Services Facility at NASA-JSC. The total mass of the air-bearing cart is 62 kg (137 lb). Compliant coupling is simulated by 2.33 mm (0.090 in) diameter wire arranged as four cantilevered beam elements. The effective stiffness of the springs is 19.34 N/m.

It was found that the total damping in the system, including auxiliary spring elements used to reduce rotation and lateral translation of the mass and the friction in the LVDT and its pulleys, was greater than the force exerted by the air jets. The free vibration of the mass is shown in figure 7. The natural frequency of the system is 0.089 Hz, and the damping factor is 0.084. The air jet identified in the previous section produces an equivalent damping factor of 0.0282. Since the damping in the experimental set up is three times the damping introduced by the air jet, the effect of the control system is expected to have negligible effect on the vibratory response.

In order to demonstrate that the control system, the following parameters are used. The air jet diameter is increased to 2.38 mm (3/32 in) and the pressure drop across the jet is increased to 8.28×10^4 N/m^2 (12.0 lb/in^2).

The vibrational response of the mass is shown in figure 3. It is seen that the air jet produces an equivalent damping, which increases the damping of the system by 0.015. In this plot, the jet, when activated, was pulsed for 0.5 second before reset. The analytical model is amended to reflect the modified jet parameters. Figure 9 shows the estimated free vibration response of the test mass. This plot correlates well with the measured free vibration shown in Figure 7. Figure 10 is the estimated response with the air-jet controller on. Again, the estimated response compares favorably with the experimental plot of Figure 8. It is found that the estimated effect of the air jet controller is strongly dependent on pulse time. The plot in figure 10 is obtained with a jet pulse of 0.1 second.

This is much less than the 0.5 second pulse set on the timer of the experimental controller. The discrepancy indicates that the air jet does not go to full flow at the instant that the solenoid is activated.

Conclusion and Recommendations

An experimental facility incorporating air jet active vibration control has been fabricated. The facility has been used to show that the air jet controller effectively damps oscillations. An analytical model has been developed which estimates the effect of the air jet controller. The model can be used as a design tool to quantify parameters such as pressure drop, flow rate and net acceleration of the mass under combined air jet and spring-transmitted excitation. The model has shown that the solenoid dynamics limit the thrust produced by the jet to 20 percent of the thrust from an ideal valve which produces full flow when activated.

Continued work in this project will be in (1) sensor development, (2) extension to general plane motion control, and (3) model development.

(1) Sensor development. The current LVDT will be replaced by a non-contacting probe, such as accelerometer, ultrasonic tracker, or laser tracker. Such a transducer eliminates the need for pulleys and thereby reduces system friction. The transducer also permits extension of the system to general plane motion with both translation and rotation.

(2) Extension to general plane motion. The control system will be expanded to three-degrees-of-freedom. Sensors will be developed which respond both to translation and to rotation of the mass. The analog control circuitry and the air jet configuration will also be modified.

(3) Model development. The analytical model will first be refined to resolve the discrepancy between measured and estimated jet thrust noted in the previous section. The model will then be expanded to the general plane motion case. The modified model will be used as a tool in the design of the experiment.

C- 3

It is expected that the same control algorithm which is obtained for plane motion control will be applicable to the more general six-degree-of-freedom application. Thus, the control system developed in the laboratory can be adapted for use in the Space Station.
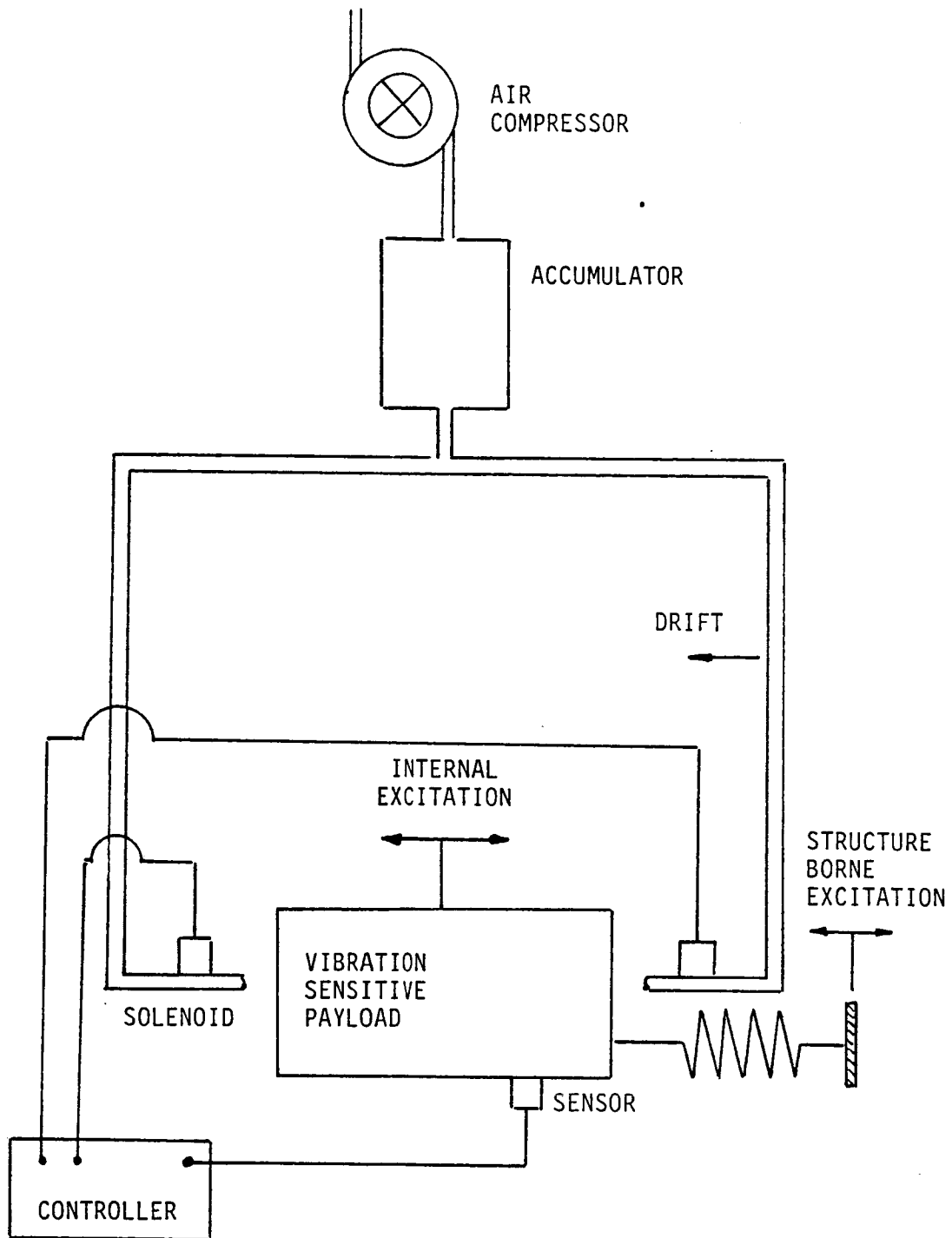
Acknowledgments

FIGURE 1. SCHEMATIC REPRESENTATION OF FEEDBACK CONTROLLED SYSTEM
USING COMPRESSED AIR AS FORCE PRODUCER
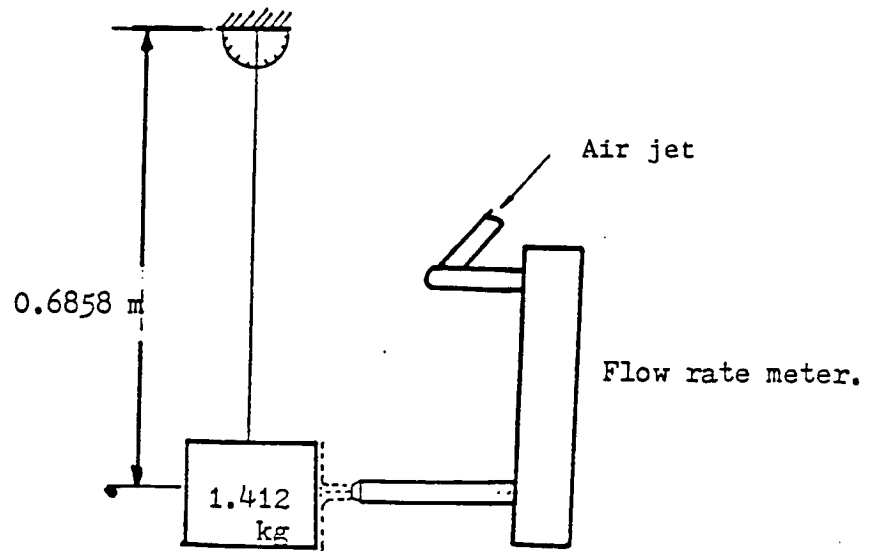
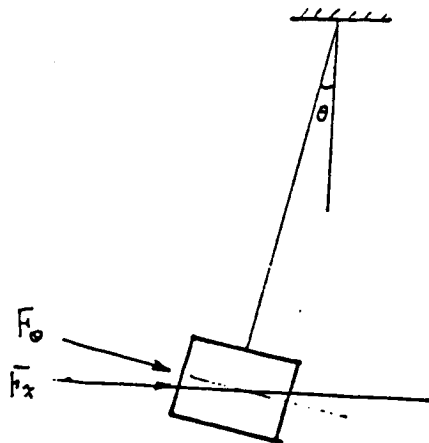Fig. 2a. Schematic of experiment.

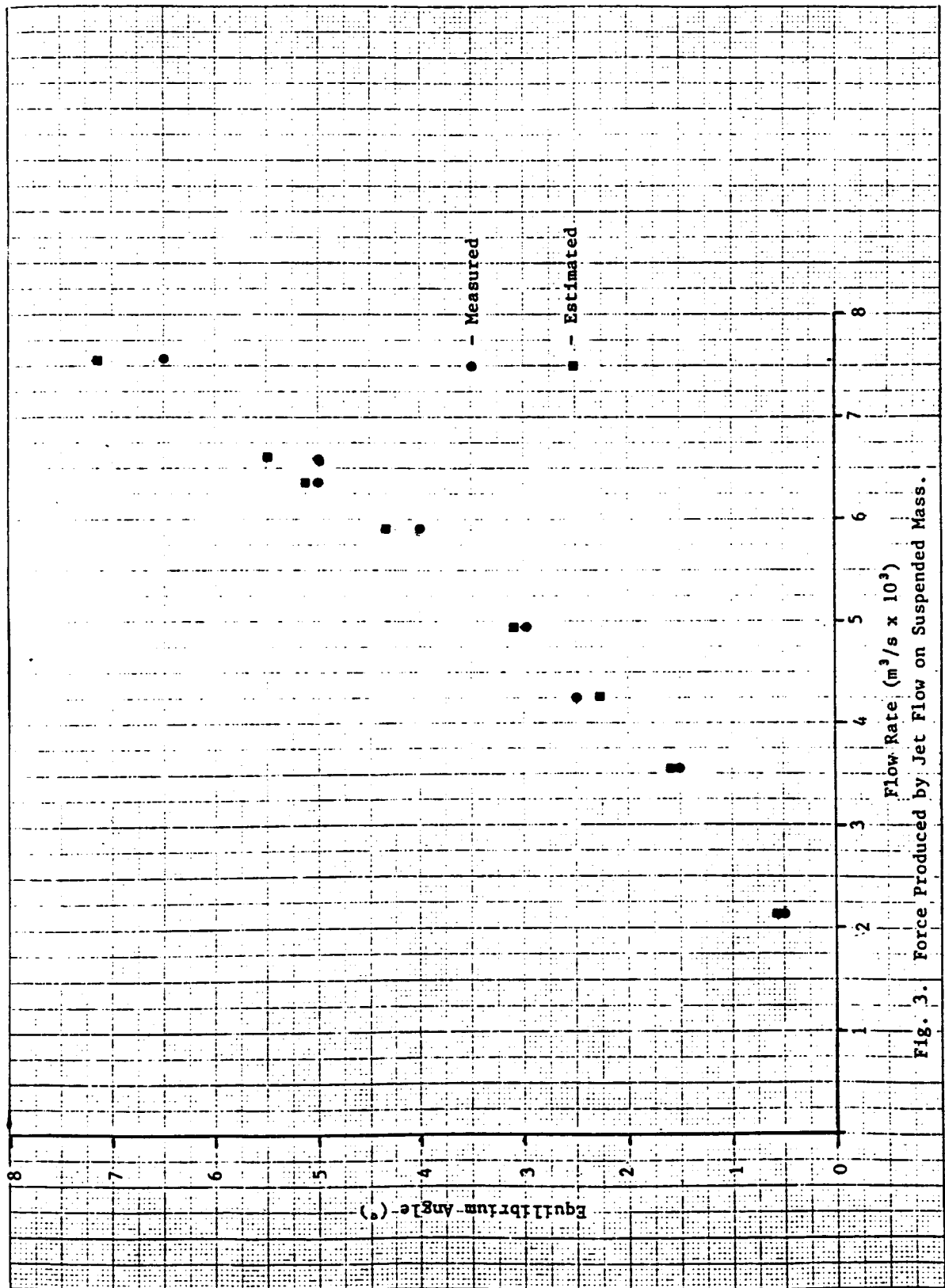Fig. 2b. Equilibrium of Pendulum due to
Impinging Air Flow.
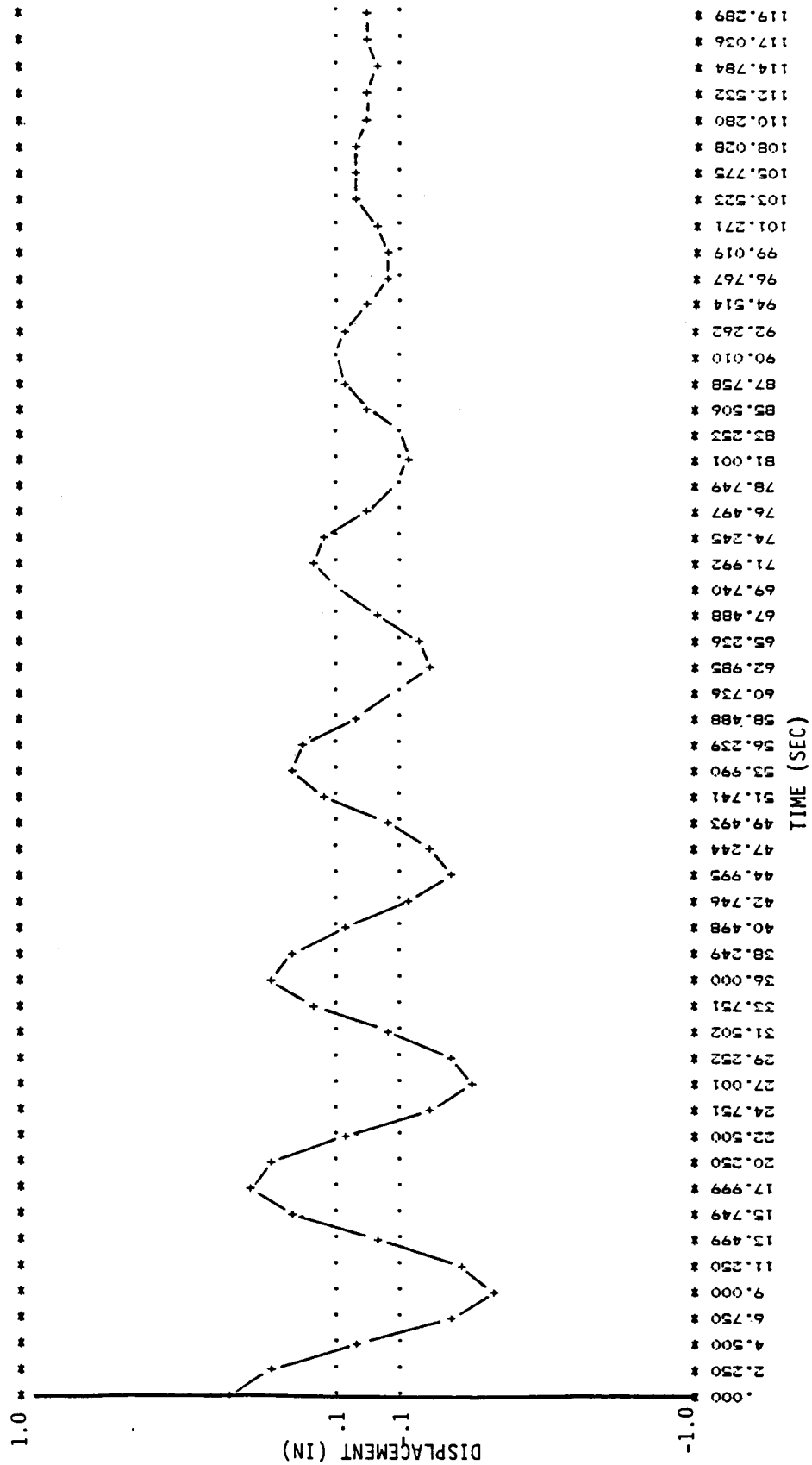
Fig. 3. Force Produced by Jet Flow on Suspended Mass.

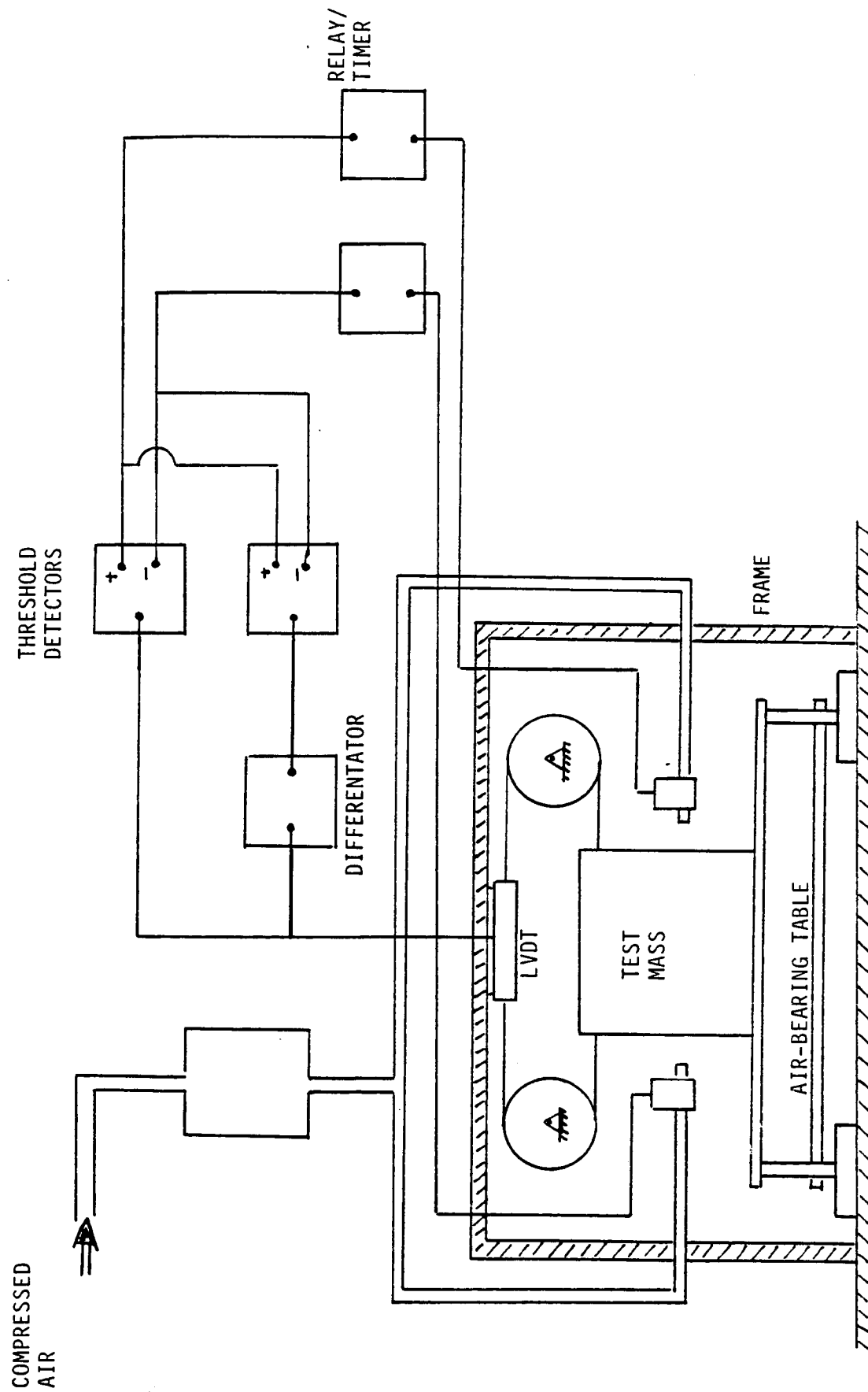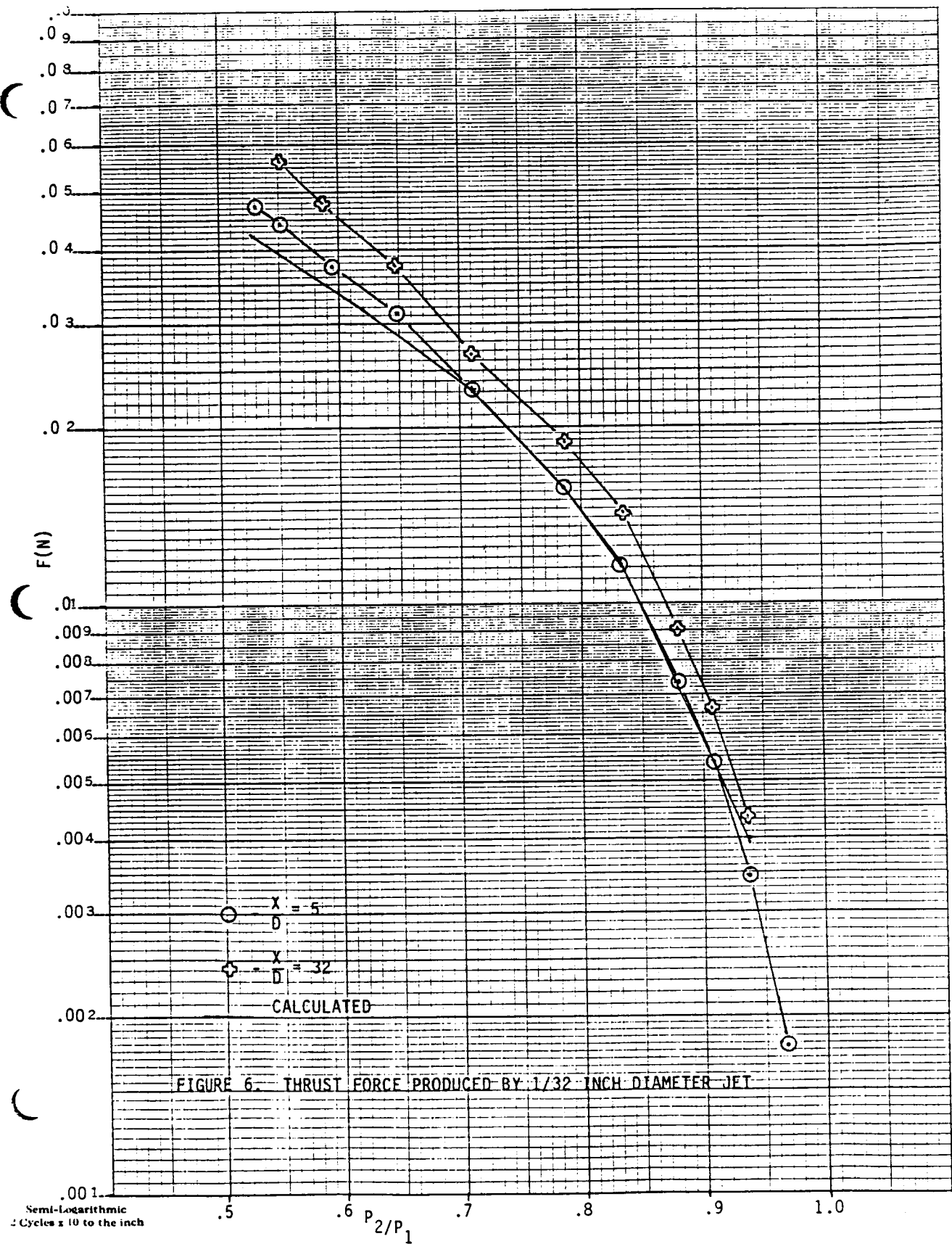FIGURE 4. EXPECTED RESPONSE OF SPRING-MASS SYSTEM WITH AIR-JET CONTROL

FIGURE 5. AIR JET CONTROLLER SYSTEM EXPERIMENT SET-UP
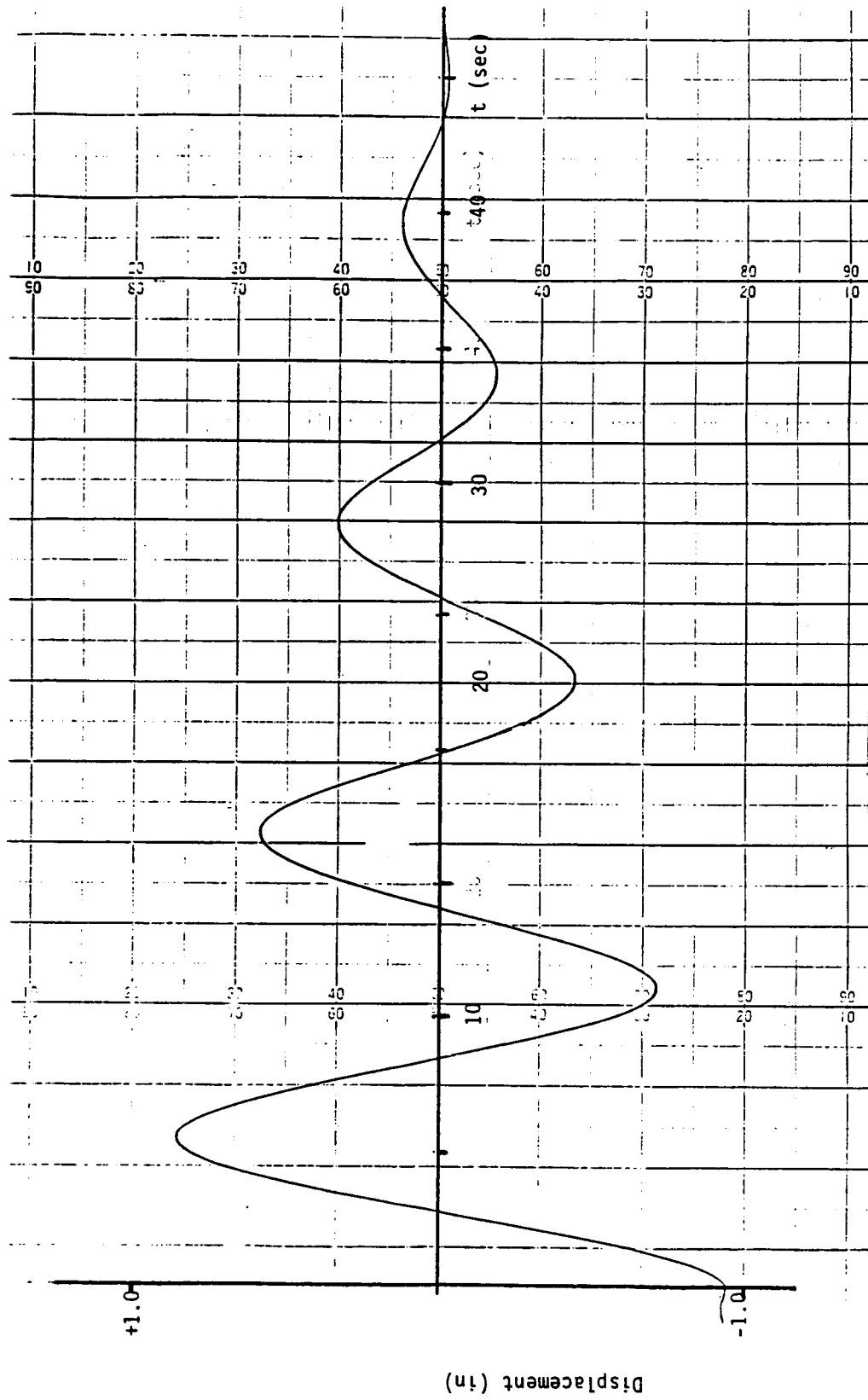
FIGURE 6. THRUST FORCE PRODUCED BY 1/32 INCH DIAMETER JET

Figure 7. Free Vibrational Response of Test Mass
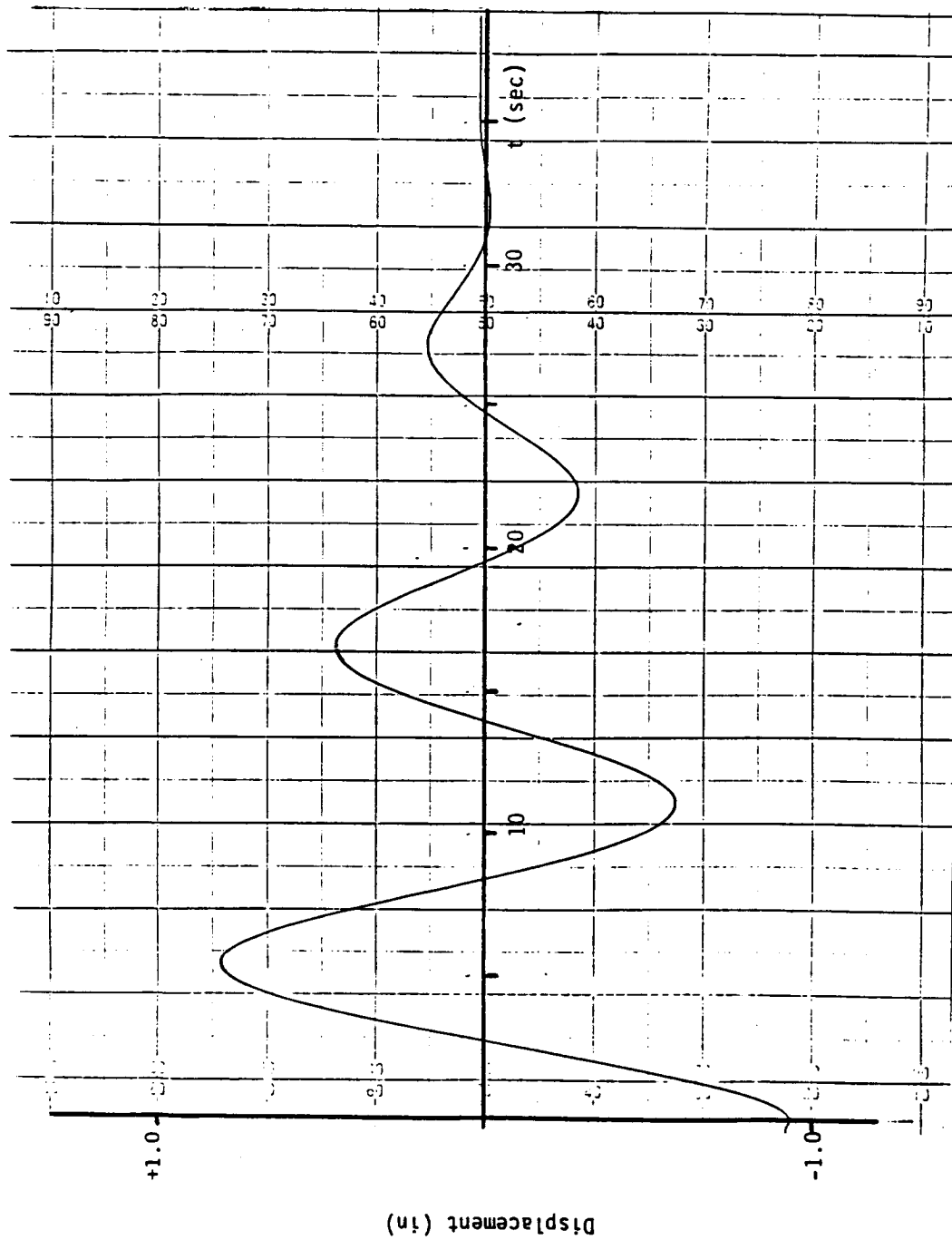
Figure 8. Test Mass with Air Jet Controller Activated
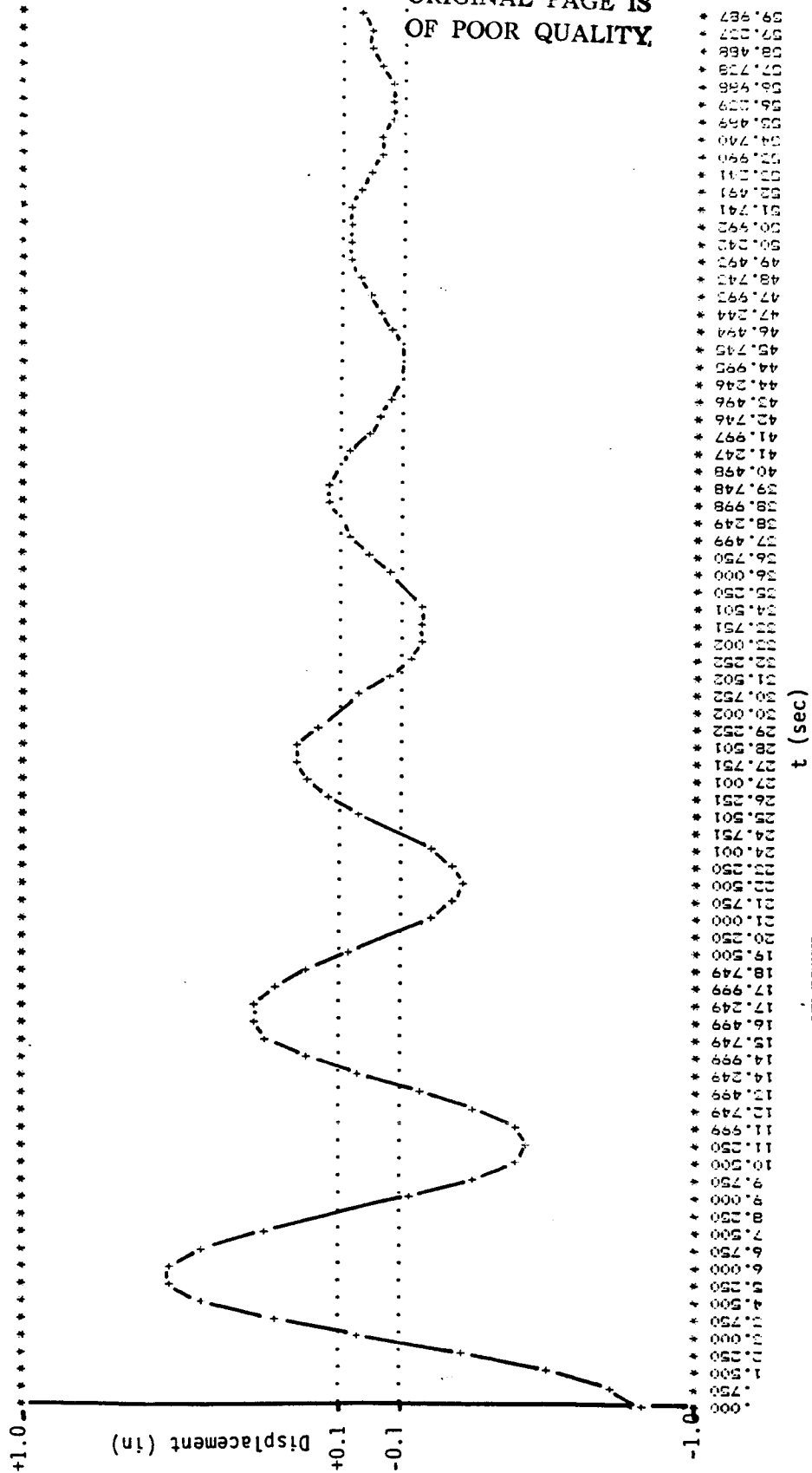(0.5 Second Pulse)

Displacement (in)

+1.0

+0.1

-0.1

-1.0

t (sec)

Figure 9. Estimated Free Vibration of Test Mass
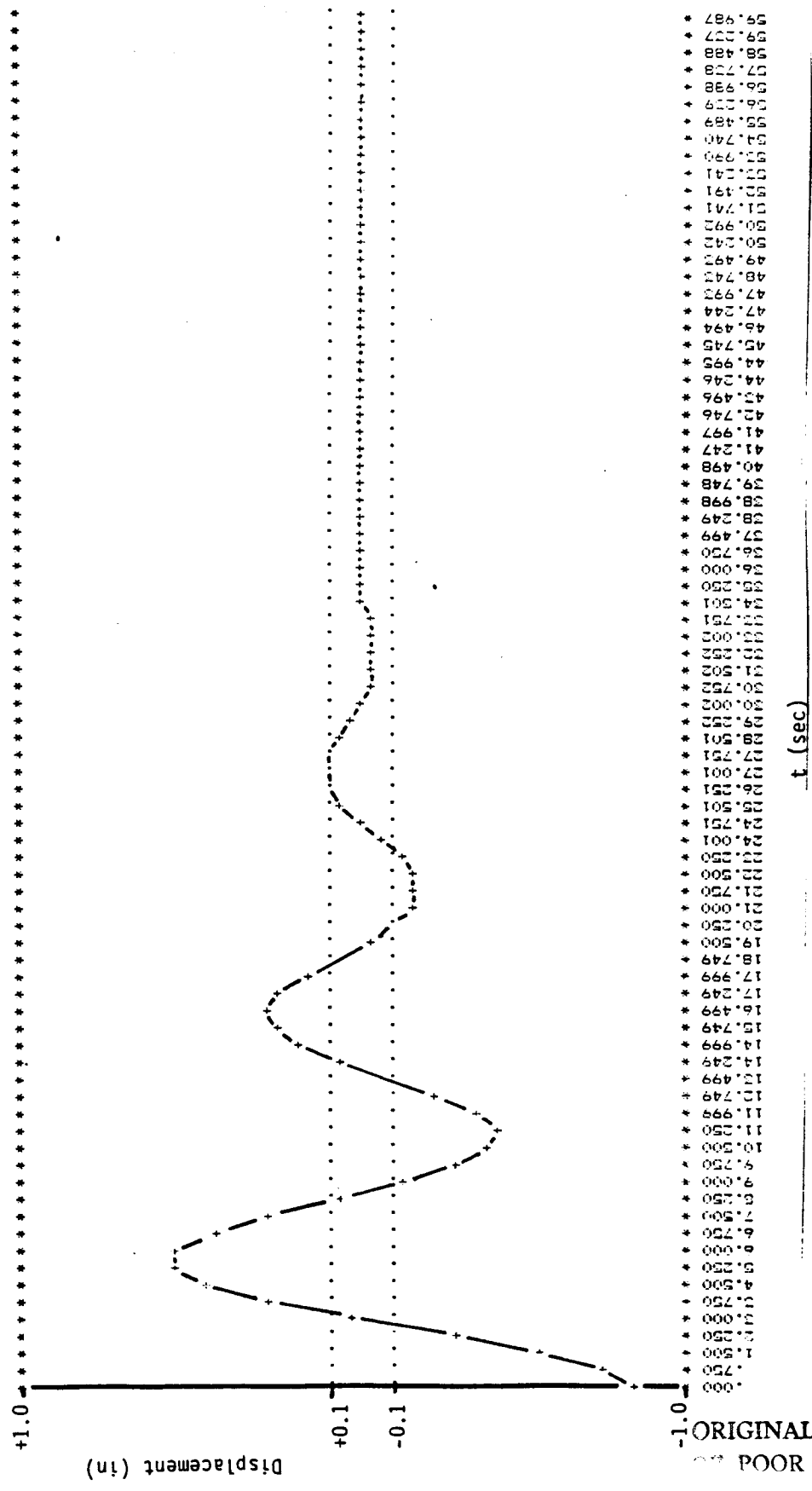
Figure 10. Estimated Vibration of Test Mass With Air Jet Control

(0.1 Second Pulse)

t (sec)

Displacement (in)

+1.0
+0.1
−0.1
−1.0

Appendix A. Derivation of Jet Flow Equations

It is assumed that the momentum flux, $\frac{dm}{dt} V_f$, for a one-dimensional,

isothermal jet is constant downstream of the jet, and equal to the momentum flux at the jet exit, M_0. For a round jet:

$$M_0 = \rho \pi U_0^2 R^2 \qquad\qquad (A\text{-}1)$$

where:

ρ = fluid density

R = jet radius

U_0 = fluid velocity at the jet exit (assumed uniform)

Downstream of the jet exit, the jet widens, the centerline velocity decreases and the velocity profile across the jet takes on a Gaussian distribution [5]:

$$\frac{U(y)}{U_\mathcal{L}} = e^{-K\eta^2} \qquad\qquad (A\text{-}2)$$

where:

$U_\mathcal{L}$ = centerline velocity at downstream station, x

$$\eta = \frac{y}{x + x_0}$$

y = radial distance from the centerline

x_0 = virtual origin of the jet, assumed to be at the jet exit.

K = constant

The constant, K, is empirically derived from the jet widening rate, which forms an included angle of 9.8° to the radial location at which the flow velocity is 0.5 the centerline velocity [6]. The value of K is 94.

In the fully developed region, beyond approximately 8 diameters downstream of the jet, the centerline velocity decreases at [6]:

$$\frac{U_{\mathcal{L}}}{U_o} = \frac{6.2}{(X/D)} \tag{A-3}$$

where:

D = jet exit diameter

As the jet widens and the flow speed decreases, air surrounding the jet is entrained, such that the momentum flux remains constant.

Using the relationships obtained above in the mass flow rate equation:

$$\frac{dm}{dt} = \int_o^{\infty} 2\pi \rho y \, U(y) \, dy$$

it is found that, for an isothermal jet:

$$\frac{dm}{dt} = 0.234 \times (M_o \rho)^{1/2} \tag{A-4}$$

Since the momentum flux is constant and equal to M_o, the equivalent uniform flow speed V_f can be calculated from

$$V_F = \frac{Mo}{\frac{dm}{dt}} \tag{A-5}$$

Reference

1. Garriott, O. K., and Debra, D. B., "A Simple Microgravity Table for the Orbiter or Space Station", NASA Johnson Space Center, Houston, Texas.

2. McDonnel-Douglas Technical report on Microgravity Vibration in the Space Station, MDTSCO document.

3. Dailey, J. W., and Harleman, D. R., Fluid Dynamics, Addison-Wesley, Reading, Mass., 1966.

4. Ogata, K., System Dynamics, Prentice-Hall, Englewood, Chiffs, NJ., 1978.

5. Gilbride, J. F., "Modified Shielding Jet Model for Twin-Jet Shielding Analysis", M. S. Thesis, Texas A&M University, August 1983.

6. Chen, C. J., and Rodi, W., Vertical Turbulent Buoyant Jets, A Review of Experimental Data, Pergamon Press, New York, NY, 1980.

1986   **N 8 7 - 2 6 7 0 1**

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center
University of Houston

TRAINING FOR LONG DURATION SPACE MISSIONS

Prepared by: Joseph H. Goldberg, Ph.D.

Academic Rank: Assistant Professor

University & Department: The Pennsylvania State University, University Park, Pennsylvania

                                Department of Industrial & Management Systems Engineering

NASA/JSC

Directorate: Engineering

Division: Advanced Programs Office

Branch: Systems Definition

JSC Colleague: John W. Alred

Date: July 23, 1986

Contract Number: NGT-44-005-803 (University of Houston)

TRAINING FOR LONG DURATION SPACE MISSIONS

Joseph H. Goldberg
Assistant Professor
Department of Industrial & Management Systems Engineering
The Pennsylvania State University
207 Hammond Building
University Park, PA 16802

The successful completion of an extended duration manned mission to Mars will require renewed research effort in the areas of crew training and skill retention techniques. The current estimate of in-flight transit time is about nine months each way, with a six month surface visit, an order of magnitude beyond previous U.S. space missions. Concerns arise when considering the level of skill retention required for highly critical, one-time operations such as an emergency procedure or a Mars orbit injection.

The objectives of this research project were to review the factors responsible for the level of complex skill retention, to suggest optimal ways of refreshing degraded skills, and to outline a conceptual crew training design for a Mars mission.

Currently proposed crew activities during a Mars mission were reviewed to identify the spectrum of skills which must be retained over a long time period. Skill retention literature was reviewed, to identify those factors which must be considered in deciding when and which tasks need retraining. Task, training, and retention interval factors were identified. These factors were then interpreted in light of the current state of spaceflight and adaptive training systems. Finally, the retention factors formed the basis for a conceptual design of Mars mission training requirements.

TABLE OF CONTENTS

<u>TRAINING FOR LONG DURATION SPACE MISSIONS</u>

It is human nature to forget highly learned information. Over time, psychomotor skills that may have been overlearned degrade into awkward movements at a later time, while ordered sequences and events rapidly become disordered. The study of human skill retention and degradation has been ongoing for many decades; useful information exists, but a comprehensive model of skill retention as a function of independent task and individual factors must still be developed.

This paper considers these skill retention factors in light of a long-duration spaceflight, such as a manned mission to Mars. Retention of finely tuned skills and knowledge is absolutely necessary for the successful completion of such a mission, yet man-machine system complexities are becomming more and more complex. These skill retention issues also have implications that go far beyond manned spaceflight. Industry must train workers, often for long periods of time and with concomitant losses in productivity. Colleges and universities are also in the business of training individuals with skills and knowledge for long-term retention. Clearly, accurate long-term retention of skills and knowledge is important for productivity and safety within the entire society.

This paper is divided into three major sections in its examination of long duration skill retention for manned spaceflight. (1) Currently proposed crew activities in a manned mission to Mars are reviewed. This information gave a concrete focus to the skill retention issues described here, and allowed bounds to be placed upon the duration and nature of training and retention. (2) Recent psychological and Human Factors literature (within the past 30 years) on factors influencing long-term retention was reviewed, concentrating on results and conclusions, rather than a critique of methodology. The purpose of this section was to provide a framework by which retention factors could be studied, and to subsequently provide the necessary framework for a future model of skill retention time and quality. (3) A section on training for space missions was included. The intent of this section was to provide a foundation from which longer duration retention training could build, and to outline the required training elements for an advanced, long duration Mars mission. While much of such a conceptual design is merely an exercise in futurist guesswork, an attempt was made to logically build on the concepts presented in earlier sections of this report. A description of the research yet needed to develop a working skill retention model was also included. Such a model would greatly aid the conceptual design of Mars and other long duration missions, as well as industrial job design. Information for Sections 1 and 3, where not otherwise noted, was obtained from personal communication and experience at the Johnson Space Center, Houston, Texas. The author wishes to acknowledge Jack James, Andy Petro, and John Alred for their valuable assistance.

1. <u>MARS MISSION ACTIVITIES</u>

A current scenario for an interplanetary mission to Mars includes several distinct phases:

1. Earth lift-off to low Earth orbit (possible at space station)
2. Taxi-transfer from space station to orbiting interplanetary vehicle
3. Transit to Mars
4. Docking with second orbital transfer vehicle at Mars

5. Land on surface
6. Reverse sequence for return mission

The time involved for such a mission is on the order of 5 to 9 months each way, depending on orbits, trajectories, etc., with a 6 month or longer surface visit. Using current and near-term technology, the transit times should not appreciably differ from this estimate, but the time spent on the surface could dramatically change with the addition of a permanent manned Martian base.

1.1. CREW ACTIVITIES

The activities to be carried out in a Mars mission are as varied as those in everyday life. They may be broken down into four areas, as shown below. Sources of information on crew activities and events include Oberg (1982), Oberg and Oberg (1986), Joels (1985), Conners et al. (1985), and National Commision on Space (1986).

Spacecraft Control and Maintenance. These critical activities will insure that mission success and crew safety will not be comprimised. Crew members, armed with automated equipment and extensive computer programs, must serve as diagnosticians, continuously asking new questions about the status of equipment. As discussed by others, maintenance will range from simple modular replacement of Lithium cannisters, to large-scale reconstruction or evacuation of space vehicle apparatus. Both dexterity and decision ability will be required.

Scientific Study. Much investigation will continue to be performed in space. Many accounts have indicated an even broader range of research topics than in previous space missions. These will include not only physical and hard science topics, but will be expanded to social and behavioral science issues including space habitability, behavioral interaction, and group power structures.

Crew Health Maintenance. The health of the crew will include both regular physiological and psychological screening. Many innovative diagnostic and treatment procedures will be developed for long duration space missions, based on prior space station experience. Intelligent computer programs, in the form of expert systems, will likely be extensively used as guides for diagnosing and studying new forms of illnesses.

Recreation. Many have stressed the importance of recreation in an interplanetary mission (e.g., Fraser, 1968). Alternative cognitive and physical activites during off-duty time will be important to maintaining a healthy crew.

1.2. FACTORS IN A LONG DURATION MISSION

Skill training and retention requirements for a Mars mission will necessarily differ from that required by all previous missions, as determined by criticality and duration of events. Those skills that quickly degrade must be refreshed often or continuously, while better retained skills need only be refreshed periodically. A training program for this mission must consider several factors that are unique to a manned interplanetary journey, as listed below.

Skill Retention Duration. The required skill retention interval, between training and actual performance, may be 6 months or more. This is an order of magnitude greater than previous U.S. missions, and presents many unknowns for complex skill and procedures retention. Special on-board refresher training will be required for spme of these degraded skills.

Crew Autonomy. As one-way communication lags of 10 to 30 minutes will be encountered near Mars, Earth-based mission control will be of little use. Instead, ground-control will serve as an independent opinion source and coach for an autonomous crew. The crew of 6 to 8 will function as a team, with each member contributing complementary expertise. Crew training thus must focus on enhancing those traits that increase this autonomy, and counter the negative effects of group thinking.

Crew Confinement. The adverse effects of long-term confinement must be well understood before undertaking this mission. Training for long-term confinement must be considered, and techniques of countering confinement, such as projecting video landscapes, may be necessary. Study of analog confinement environments, such as prison or arctic stations, will aid in this definition. .

Criticality of Skills. Some required skills, such as orbital docking, will have a criticality beyond all other skills. Many of these will be performed only once or twice in a mission, after a long no-practice duration. The effects of real and perceived skill criticality on performance and training must be understood before undertaking a Mars mission.

Automation. Extensive use of artificial intelligence and automated sensing and diagnosing apparatus will be used for routine spacecraft control and maintenance. The crew will be responsible for monitoring this equipment, and factors determining crew monitoring or vigilance performance must be understood. A useful human-machine allocation model must be developed, and training for this will be required. NASA has already taken a first step in defining this model (von Tiesenhausen, 1982).

Workload. The effects of mental and physical workload must be modeled before initiating a long mission, to allow a constant performance level within an autonomous crew. The choice of how many crew members to allocate to tasks should be determined via a generic workload modeling computer program.

Environment. The adverse effects of vibration, noise, radiation, ion concentrations, and carbon dioxide are among the many environmental factors whose effects will be felt over the entire mission. The effects of these factors on health and skill retention must be considered in the design of the Mars mission.

These important factors must all be considered when designing a training program for a long duration mission. While shorter duration mission crews have tolerated and even performed well under some of these factors, their effects will be exacerbated by long-term confinement. Since a Mars mission is an order of magnitude beyond current missions in duration and complexity, its training program cannot be evolutionarily developed. Instead, a rethinking of training is required; a model specifying training needs by type of skill and degradation level must be developed. The purpose of this paper is to take an initial step towards such a model, by indicating those factors that affect skill retention, and thus training requirements.

2. SKILL RETENTION

The duration and quality of skill retention should necessarily determine the training requirements of a long duration space mission. Skills that quickly degrade must often be refreshed, whereas better retained skills may be neglected for a longer time. Before considering this literature, several qualifications must be made, however. (1) Reports of studies in this area are often not readily obtainable. This may be due to the fact that much of the training research has been conducted in the private and military sectors, which have little impetus to publish in widely distributed publications. Also, much of this research is very task specific, and investigators may have have felt that their research would have low utility outside their immediate scope. (2) The major retention factors are covered below as discrete topics, but all are intimately intertwined and confounded. Differences in the length of a post-training retention interval, for example, are confounded with the type and duration of initial training. Conclusions drawn here must clearly be interpreted with a great deal of caution. To gain a better understanding of these factors, however, they are discussed separately, ignoring conjoint and interactive effects. (3) In some cases, conclusions were necessarily drawn from very few studies, clearly scientifically inappropriate. This was pragmatically done to at least provide a direction for future research needs and developments.

Naylor and Briggs (1961) reviewed over 60 years of literature, and created the first categorization of retention-influencing variables. (1) Task variables included the procedural/tracking task dichotomy introduced below. They raised the important issue that the difficulty and organization of a task is likely responsible for observed retention differences. (2) training variables included three subclasses of factors: the amount of initial training, distribution of training sessions, and transfer effects from other tasks. (3) Retention interval variables included those factors present within this period. (4) Recall variables consisted of other retention-influencing factors, such as the training fidelity, or the presence of any warmup activity prior to retention testing. The present review drew heavily on this work, and extended their factor categories. A subsequent review (Gardlin and Sitterley, 1972) covered many skill retention studies, under contract to NASA. These investigators provided annotated reviews of many studies that were directly applicable to the piloting of space vehicles. The present review also drew on this paper, but was broader in coverage.

Selected skill retention studies cited below are summarized in Table 1, which presents the following information: (1) Investigator(s), (2) Retention: time interval between end of training and initial retest, (3) Task: type of performance task. P: procedural (discrete), T: tracking and control (continuous), (4) Indp. Var.: independent or manipulated variable(s); D: duration of training, R: retention interval, S: structure of training, F: fidelity of training, O: organization of task, RR: retention interpolated activity or rehearsal, (5) Task Description, (6) Cmplxty: complexity of the task(s), subjectively estimated by the number and type of simultanous activities that had to be performed, (7) Training: method of training; duration or criterion, (8) #Ss: number of subjects tested across entire study, (9) Exper: subjective subject experience at task; all subjects were inexperienced in abstract tracking tasks, whereas some aircraft control studies utilized experienced pilots; I: inexperienced, E: experienced, (10) Dependent Var.: dependent or measured performance variable(s)

The results of these studies, referred to throughout this report, are summarized in Figures 1, 2, and 3. Manipulations of retention interval, training duration, and training organization are shown, and every attempt was made to combine similar studies with identical dependent parameters into one figure. Figure 1 shows performance in procedural tasks, or those requiring cognitive control or sequencing over many procedural steps. Note that three time scales, to allow sufficient resolution, were used on the retention time axes: 0-24 months, 0-6 months, and 0-4 weeks. Dependent variables here included both errors and time to complete procedures. Figure 2 shows performance in simple tracking tasks over the same retention intervals as the procedural tasks. Dependent measures here included integrated error in volts, inches, or arbitrary numbers, measuring the deviation between a target and one's ability to follow it. Other measures included the acqusition time to capture a target, or the percentage of total time on a target. All of these parameters generally required some form of continuous sampling by the experimental apparatus. Figure 3 also shows continuous tracking, but only for studies which presented much more complex flight control tasks. These experiments often used open- or closed-loop simulators of airplanes or space vehicles. All retention interval axes here were 0-6 months. Dependent measures usually consisted of a large collection of parameters, of which a subset was chosen, such as altitude error from a preset flight path.

Table 1. Skill Retention Studies

Investigator(s)	Retention	Task	Indp.Var	Task Description	Cmplxty	Training	#Ss	Exper	Dependent Var.
Neumann and Ammons (1957)	1.75, 12 months	P	R	Switch setting for abstract task; 16 steps	Low	Actual task/ to criterion: 2 error free trials	40	1	Number of correct switch settings
Ammons et al (1958); Exp 1	1,6,12,24 months	P	D,R	Switch and control setting for an abstract task; 17 steps	Low	Actual task/ 5 vs. 30 trials	538	1	performance time
Ammons et al (1958); Exp 2	1,6,12,24 months	T	D,R	Pedal and stick control of model airplane	Medium	Actual task/ 1 vs. 8 hours	465	1	Mean time on target
Mengelkoch et al (1960; 1971)	4 months	T,P	D	Partial closed-loop control of airplane simulator; 50 min. flight	High	Classroom; Simulator/ 9.0 to 13.5 hours	26	1	Procedural errors; Flight control error measures
Fleishman and Parker (1962)	1,5,9,14 months	T	S,R	Partial closed-loop control of aircraft simulator; tracking dot on oscilloscope	Medium	Self-study; simulator/ 6 hours	70	1	Integrated tracking error
Naylor et al (1962; 1968)	1 month	T,P	D,O	Procedural switch setting (9 steps) & 3-D tracking task	Medium	Simulation; part- and full-task/ 2 vs. 3 weeks	128	1	Integrated tracking error; Procedural response time and errors
Hammerton (1963)	6 months	T	D	Second-order (acceleration) tracking of dot on CRT	Medium	Actual task/ to criterion	18	1	Target acquisition time

Table 1 (cont). Skill Retention Studies

Investigator(s)	Retention	Task	Indp.Var	Task Description	Cmplxty	Training	#Ss	Exper	Dependent Var.
Trumbo et al (1965a)	1, 5 months	T	D,O,R	Tracking line on oscilloscope	Low	Actual task/ 50 vs. 100 trials	250	I	Integrated tracking error; temporal and spatial accuracy
Trumbo et al (1965b)	1 month	T	D	Tracking line on oscilloscope	Low	Actual task/ 115 vs. 130 trials	120	I	Integrated tracking error; other tracking indeces
Grodsky et al (1964; 1966)	1, 2 months	T,P	R	Full-scale simulation of 7 day lunar landing mission	Very high	Simulator	3	E	Procedural errors; flight control error
Cotterman and Wood (1967)	4,8,9,13 weeks	T,P	R	Full closed-loop control and landing of lunar excursion module	Very high	Simulator/ 6 weeks	12	E	Probability of successful performance with respect to criterion
Swink et al (1967)	3, 5 months	T	S,O,D,R	Tracking line on oscilloscope	Low	Actual task/ to varying criterions	120	I	Integrated tracking error; temporal index
Youngling et al (1968)	1, 3, 6.7 months	T	D,R,S	Closed-loop control of space vehicle flight simulator	Very high	Simulator/ 60 vs. 120 trials	96	I	Mean time on tracked target
Grimsley (1969)	1 month	P	F	Setting switches and controls on panel; 92 steps	Medium	Actual task, varying fidelity/ to criterion	60	E	Number correct control settings; training time; time to retrain

Table 1 (cont). Skill Retention Studies

| Investigator(s)	Retention	Task	Indp.Var	Task Description	Cmplxty	Training	#Ss	Exper	Dependent Var.
Sitterley and Berge (1972)	1,2,3,4,6 months	T,P	R,RR	Closed-loop control of space vehicle simulator; lift-off to orbit insertion	Very high	Simulator; classroom/ ave. 34 flights	45	I	Flight control error measures
Sitterley et al (1972)	4 months	T,P	RR	Closed-loop approach and landing of Shuttle simulator	Very high	Simulator; classroom/ to criterion (ave 10 hrs)	15	E	32 Flight control error measures; procedural completion time
Sitterley (1974)	4 months	T,P	F	Closed-loop approach, landing; Shuttle simulator; 7 min. flight	Very high	Self-study; Classroom; Simulator/ 8.1 hours	5	E	Integrated flight performance
Shields et al (1979)	4, ,12 months	P	R	20 common soldier tasks; e.g., donning gas mask	Very high	Army basic training	523	E	% go/no go task performance
Schendel and Hagman (1980)	2 months	P	D,S	Disassembly and assembly of M60 machine gun	Medium	Actual task/ to criterion	26	E	Procedural errors
Johnson (1981)	8-11 weeks	P	F,X	Setting controls for a simulated industrial process	Medium	Simulator; Photos/ to criterion	60	I	Procedural errors; performance time; device transfer
Goldberg et al (1981)	5 weeks	P	D	Boresight and zero main gun of M60A1 tank, 27 steps	Medium	Actual task/ to criterion	42	E	Procedural errors
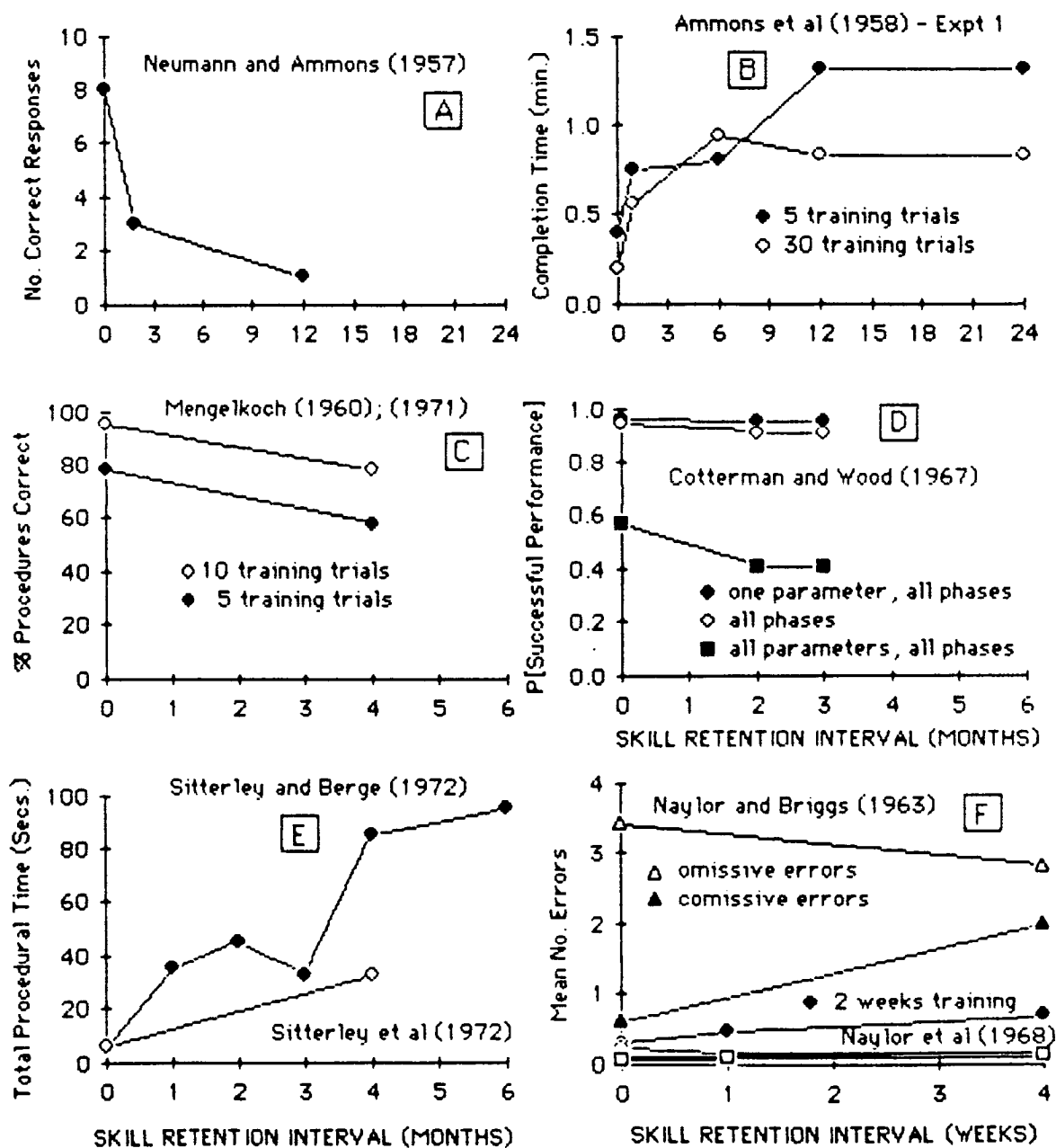
Figure 1. Procedural Skill Retention as a Function of Time Interval.
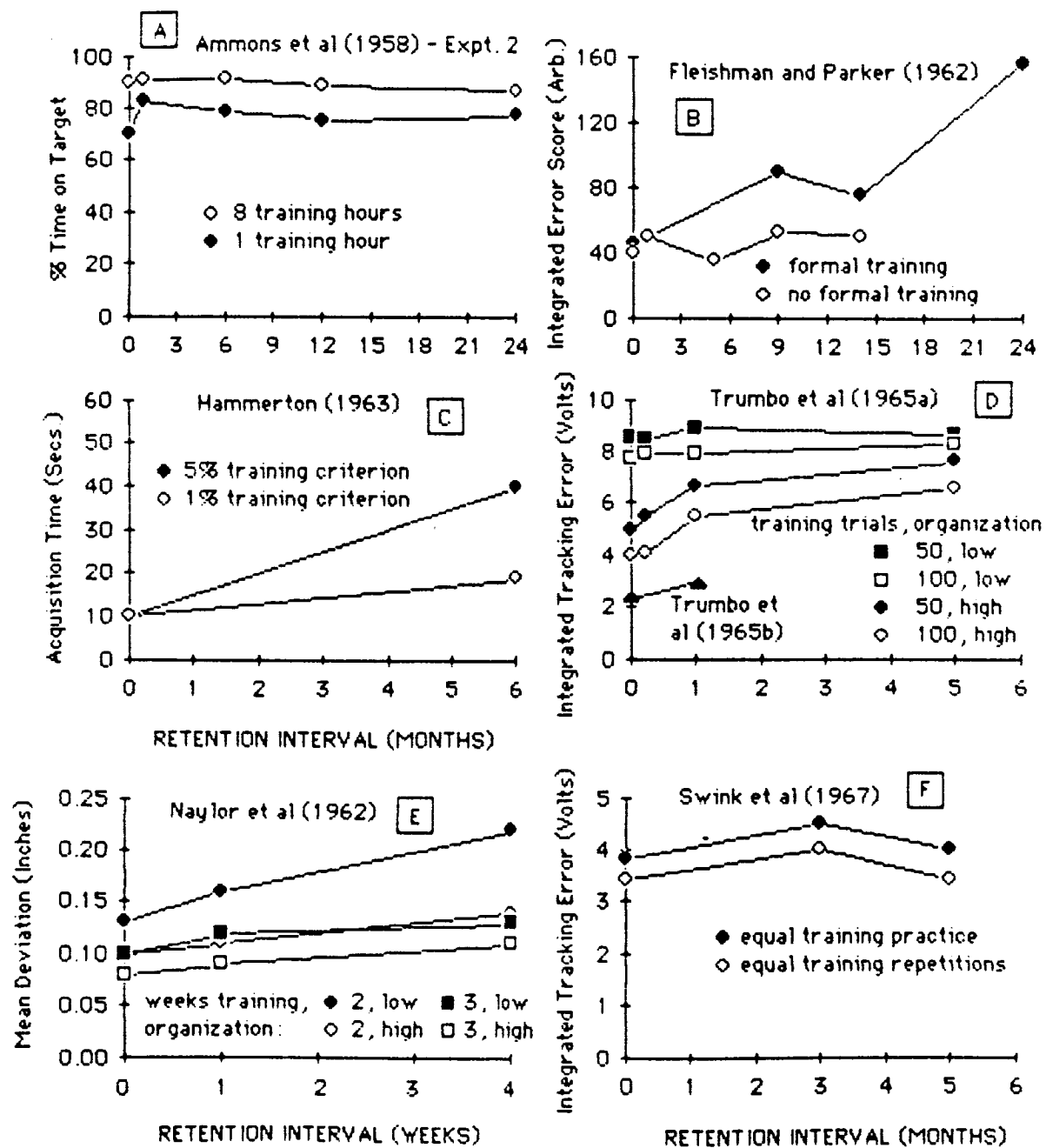Note Varying Retention Interval Axes:

Figure 2. Simple Tracking Skill as a Function of Retention Interval.
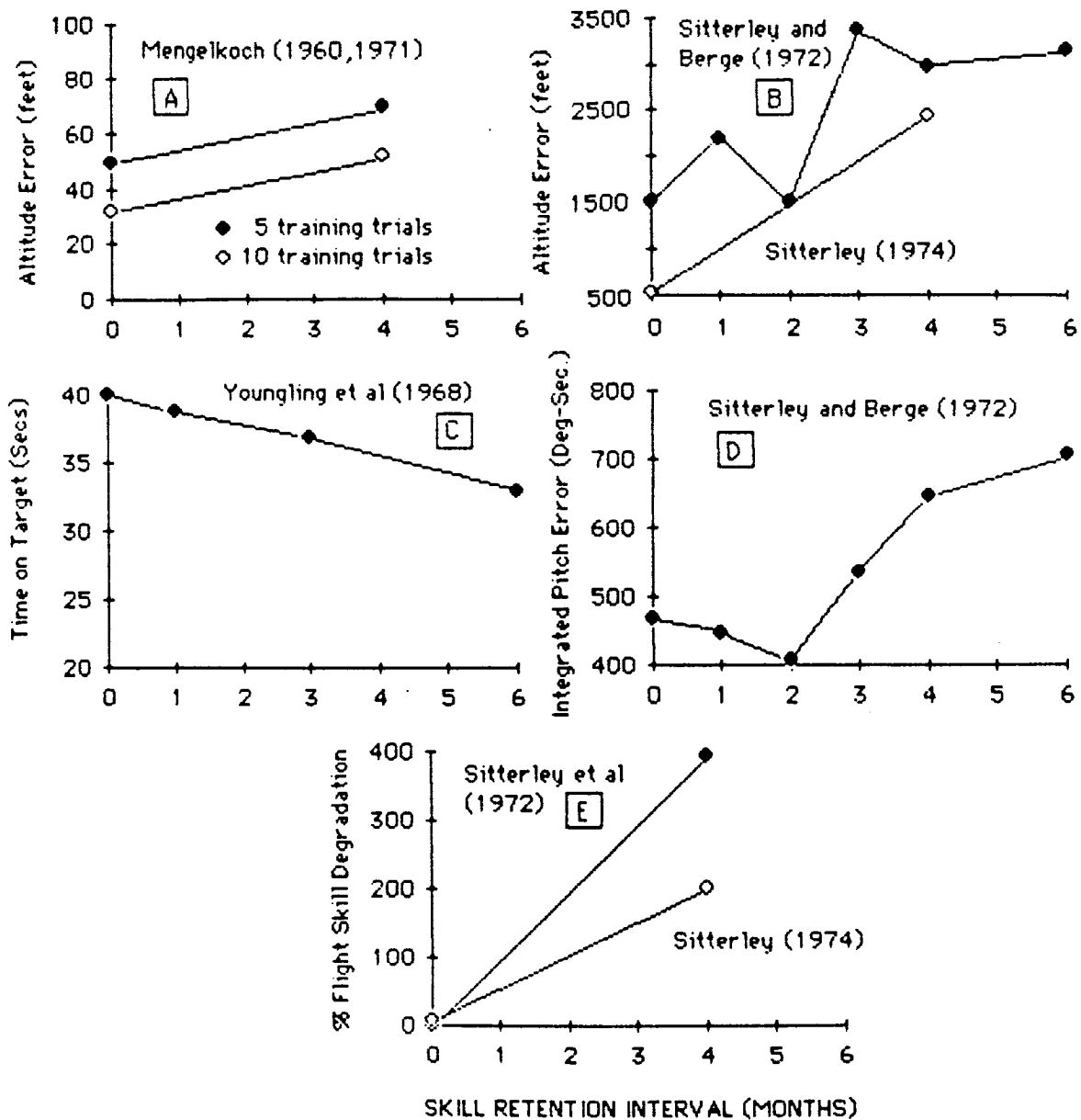Note Varying Retention Time Axes.

Figure 3. Complex Flight Control Skill as a Function of Retention Time Interval

2.1. TASK FACTORS

Those factors affecting skill retention that are direct properties of a task are considered in this section. Clear trends that appear in these factors are valuable in the design and evaluation of task training, as those factors responsible for a large amount of skill degradation are good candidates for potential elimination or control. Presence of a large number of these critical factors can point to tasks requiring frequent refresher training. Great care must be used when evaluating tasks containing more than one of these factors, as interacting factors have rarely been investigated in a controlled manner. Anything other than qualitative comparisons across studies are dangerous, due to countless numbers of uncontrolled factors. Rather, generalizations should be drawn by first noting within-study conclusions, then qualitatively comparing these across many studies.

2.1.1. Type of Task

Since some of the earliest skill retention research, a distinction has been made between overall types of tasks. Procedural tasks are those requiring discrete, ordered responses. Some have labelled these as cognitive tasks, referring to the large amount of non-automatic cognitive resource required when trying to recall a long sequence of task steps, while others have called them discrete tasks, for the isolated responses that are required. Examples here include checklists on aircraft or space vehicles, emergency procedures, and more abstract tasks such as setting sequences of switches in a proper order. The other class of tasks are those which require psychomotor skills for successful completion. Some have also labelled these as continuous control or tracking tasks, because discrete responses are not given. Typically, these require an individual to keep some stimulus on target, or within a specified range of conditions. Common examples here are driving a car, controlling an airplane or space vehicle, or simply manually controlling a lever so that a displayed shape remains between two points on an oscilloscope. Most real-world tasks contain an element of both of these. Early research efforts simplified these as much as possible to obtain a high degree of experimental control. These studies indicated that procedural skills degrade more quickly than operational or continuous skills. Only those studies which presented both types of tasks to subjects, measuring relative skill degradation differences, are appropriate in the present assessment.

In early studies, continuous tasks were simple tracking movements while procedural tasks consisted of sequence memorization. Ammons et al. (1958) tested over 1000 subjects on either a 17-step procedural task or a model airplane control task, over many retention intervals. Greater skill loss occurred on the procedural, cognitive task than on the motor control task over retention intervals of up to two years. Naylor et al. (1962) combined a procedural switch.setting task (9 sets of 3 switches) with a three dimensional joystick-controlled tracking task and found similar skill retention for both tasks. In this instance, however, Gardlin and Sitterley (1972) noted that the procedural task was very simple, and greater skill degradation relative to the tracking task might have otherwise been expected. Using only a tracking task, Trumbo et al. (1965a) broke down overall performcance into both temporal and spatial accuracy dependent measures. The former measured anticipations and lag time, whereas the latter measured absolute positioning accuracy. Interestingly, the temporal skill performance was lost more quickly than the spatial performance over the one week to five month retention intervals. The better subjects may have emphasized the temporal aspects of the task more than the spatial aspects, suggesting that more effort should be spent on maintaining temporal task performance.

Later studies have used more complex, flight control tracking tasks with procedural checklists, measuring performance on both. Mengelkoch et al. (1971) combined a flight control task with a procedural checklist task containing 125 discrete items. Over a four month retention interval, discrete procedural responses were more susceptible to forgetting than the continuous flight control responses. Though the procedural losses were great enough to be practically significant, the investigators were careful to qualify this conclusion. It is not possible to define equal levels of learning between the two types of dependent measures; tracking is measured as a continuous variable (e.g., altitude error), while procedural lists are measured by percentage error. In this context of real aviation training, however, more emphasis should logically be placed on the learning and retention of procedural checklists than on continuous flight control (Gardlin and Sitterley, 1972).

Sitterley and others conducted a series of skill retention studies for NASA in the early 1970's. All tasks required both active flight control and the use of procedural checklists, and utilized a complex, closed-loop space shuttle cockpit simulator. Comparing manual control with emergency procedural skill retention from one to six months, Sitterley and Berge (1972) found that procedural skills degraded much more rapidly than operational skills. Flight control skills were acceptably retained for two months, whereas procedural performance degraded after only one month; flight performance degraded by a factor of ten after an interval of four months. Sitterley et al. (1972) found similar patterns of degradation, but the procedural skill loss was not as great as that found in the previous study. The studies differed in that the present study used experienced pilots (the previous did not) and allowed warm-up techniques prior to retention testing.

The studies cited above have all had in common the performance and measurement of both continuous, tracking and discrete, procedural skills. While differing retention intervals and task complexities were used, the consensus has been that procedural skills are (1) more quickly lost, and (2) lost to a lower relative skill level than tracking skills.

The underlying factors responsible for this differential retention loss are still unknown, however. Task organization differences may be responsible for this differential (see Section 2.1.2.). Alternatively, proportionally more training may be achieved by tracking than by procedural performance in a short period of time, because of its continuous nature, as discussed in Section 2.2.1. Also, there may naturally be more transfer and practice of tracking skills in a given retention interval than is allowed for procedural skills (see Sections 2.2.3. and 2.3.2.). The fact that procedural skills degrade more rapidly and fully than operational skills may thus be an emergent property of other underlying mechanisms. Different types of parameters are also measured in these two classes of tasks. Procedural skills are measured by accuracy in following the established order of a task, while operational skills are typically measured by temporal parameters. If procedural tasks are measured by such parameters as the time required to complete a sequence of switch settings, it in effect becomes an operational task, so the true measure of task type may lie in the parameters measured, not in the actual task itself.

2.1.2. Task Organization

The actual or perceived task organization, in addition to the procedural/operational dichotomy, influences skill degradation. Procedural tasks may have less spatial and temporal organization

than tracking tasks (Gardlin and Sitterley, 1972; Trumbo et al., 1965; Swink et al., 1967; Noble et al., 1967). Unless a study manipulates an organizational variable in a highly controlled manner, the task type and organizational variable will be confounded.

Perceived organization has been most commonly manipulated by altering task predictability. Using a procedural task combined with a tracking task, Naylor et al. (1962) systematically manipulated task organization by illuminating light pairs in a predictable or an unpredictable order. The task organization had a greater influence for lesser trained conditions, in that more organization was required as less training was given. Trumbo et al. (1965a) created four differential conditions of target predictability in a tracking task by selecting tracking targets so that every one followed in spatial sequence, or every one, second, or third target was chosen randomly. At retention testing, subjects who received the most predictable target sequence retained skills better than all other levels of predictability. In fact, performance after a five month retention interval using the predictable sequence was superior when compared to performance after only one week of retention using the less predictable sequences. When compared with the predictable target organization, the less predictable tasks showed 80% to 100% more error, but little practical difference was noted between the low organizational task performances. Swink et al. (1967) had tracking targets either appear in a deterministic order on every trial, or with every fourth target randomly selected. Again, the predictable target sequence produced superior tracking performance for retention intervals of both 3 and 5 months. In another horizontal tracking task, Trumbo et al. (1967) produced three levels of target predictability, corresponding to every 4th target, 6th target, or no targets randomly assigned in a sequence of 12 target locations. All levels of predictability produced the same performance by the end of training sessions, but retention at the end of one week was greatest with the most predictable task. The lowest predictability task produced better performance than the medium predictability task by the end of the retention testing session, perhaps due to the fact that differential training had been given to these two groups of subjects; low predictability subjects received 195 training trials, while the medium predictability subjects received only 80 trials. Noble and Trumbo (1967) reviewed a number of experiments, breaking down retention loss by spatial and temporal uncertainty variables. In general, the greatest retention losses were noted in the most uncertain task conditions, and response strategies by subjects varied with the amount of task uncertainty.

Manipulations of procedural task uncertainty have also been used to alter organization. Naylor et al. (1968) manipulated the predictability of a secondary procedural task, while subjects were simultaneously performing a primary tracking task. Subjects had to depress buttons in varying orders, depending on which of several lights were illuminated at a point in time. Two levels of procedural task organization were defined: a light sequence in numerical order, and a sequence in random order. After retention intervals of 1 to 4 weeks, the well organized secondary task had lesser performance decrements than did the less organized task. In addition to decrements on the secondary task, performance on the primary tracking task was also influenced. The well organized secondary task produced superior retention on the primary tracking task after both 1 and 4 weeks, than did the more poorly organized secondary task.

These have clearly demonstrated that task organization directly influences both procedural and tracking skill retention duration. The actual task predictability was manipulated; retention may also be influenced by perceived organization within a task. While no studies have compared actual

versus perceived predictability, this is precisely the function of training.

2.1.3. Task Workload

Workload refers to the physical and cognitive effort imposed by an operational task. Earlier work in this area studied the effects of time-sharing between several visual displays, as is required when driving or flying, while later work utilized other, joint combinations of abstract and operational tasks. For many reasons, some effort has been made to develop models and measures of workload in generic tasks (e.g., Brown, 1978). Rouse (1979) has reviewed information theory, control theory, and queueing theory models of operator workload, as well as performance, physiological, and subjective workload measures. Workload is of interest due to its effect upon skill training and retention. Tasks requiring many simultaneous control movements, as in aircraft control, or tasks containing long strings of branching point decisions, can both be considered to be of high complexity and required workload. Johnson (1981) suggested that the number of steps required in a procedural task should be a determining factor in its probability of successful skill retention.

It is interesting to ask (1) whether skills for a high workload task are retained for a different time period than those for a lower workload task, and (2) whether an individual can be trained to perform a given task under higher workload conditions. The organization of a task may be considered to directly affect workload, in that the two are inversly related (see Section 2.1.2.).

The most valid method of increasing workload has been to add a secondary task on top of a measured primary task. Garvey (1960) trained subjects for 25 days on a tracking task, then added a different secondary task on three subsequent days. Inclusion of the secondary tasks greatly increased tracking error to levels above initial, unpracticed levels. Single-task, low workload training did not transfer to dual-task, higher workload tasks. Briggs and Wiener (1966) noted that higher fidelity training is required in high workload, dual-task performance, than in lower workloads. This result was generalized to flight control simulators. Rudimentary flight control, having low time-sharing requirements, may be trained on low fidelity devices, but greater workload requires a higher-fidelity simulator. Trumbo et al. (1967) combined a tracking task with a verbal number anticipation task, of varying difficulties. Addition of the secondary task again dropped tracking performance to below that at the start of training. Performance after retraining did not increase to the level shown by those not performing secondary tasks. Further, performance loss after 8 days of retention was independent of the introduction of the secondary task. Naylor et al's (1962, 1968) subjects performed in dual-task combinations of a tracking task with a switch-setting procedural task, with predictability in the procedural task manipulated. This also influenced workload, because much more attention had to be placed on procedural task performance in conditions of low predictability. For each of the two tasks, the low predictability procedural conditions produced both poorer absolute performance and poor retention after a 4 week retention interval. The amount of training was the greatest predictor of absolute performance level. Gopher and North (1977) combined a primary tracking task with a secondary digit-processing task, and manipulated training conditions. Greater performance improvements from training occurred under dual-task than under single task conditions, as if the motivation from a harder task was beneficial to learning.

The definition and measurement of mental workload is a young science, but some investigators have clearly implicated it as a factor in training and retention. Higher workload tasks, defined by large levels of cognitive time-sharing, are harder to learn and retain. Large performance improvements, due to training, have been noted, however. Relative skill retention duration between single and dual-task conditions also remains to be modeled.

2.1.4. Performance Measurement

The methodology used to assess one's performance in a task has been a continuing issue for many years. In fact, an entire issue of a journal was recently devoted to this topic (Human Factors, Volume 21(2), 1979). Measurement of procedural and operational skill following a no-practice retention interval has relied on measures of accuracy or speed. These dependent variables are then plotted as a function of time, across several experimental subject groups. Bahrick (1964) critiqued the basic skill retention curve, suggesting that observed changes are due to varying sensitivities of an overserver's perceptual/cognitive system between testing sessions, and not necessarily due to forgetting or retention differences. Converting degradation scores into Z-scores may aid in stabilizing the variance between testing sessios (Bahrick, 1965).

Single-task performance measurement may not capture the concurrent demand, time-sharing requirements of real work environments, and many have artificially combined many tasks together into multiple task batteries (e.g., Alluisi, 1967). These task batteries provide high validity, precision in measurement, and sample a broad range of abilities (Akins, 1979). However, some have argued that the batteries are unnecessarily artificial, and performance scores may be defined rather arbitrarily (Chiles, 1967; Akins, 1979).

Swezey (1979) introduced a Bayesian-oriented utilities model to determine what criterion level should be achieved at the end of training sessions for gunnery trainees. This called for a 10-step decision model, identifying components of the model and calculating utility. Other, empirically-based methods of assignment to training programs have also been presented (e.g., Savage et al., 1982).

The appropriate choice of useful performance variables and methodologies is still very much at issue, particularly in light of the fact that the degree of observed retention is dependent on which perameters are used. Some investigators have measured absolute performance, whereas others used difference scores, subtracting post-training performance from retention performance. Both types of scores are required to evaluate loss during a no-practice retention interval (Gardlin and Sitterley, 1972). Also, variance measures, in addition to means, have rarely been used as performance indicators.

2.2. TRAINING FACTORS

This section covers those factors having their primary influence on initial task training. Many have suggested that these variables are at least as influential as task variables in determining the duration of skill retention. Factors covered here include the duration of training, the distribution of initial training sessions, transfer of training, and fidelity issues. The central questions to consider concern the required amount of training (cost) to expect adequate skill retention for a

desired interval (benefit). Such issues as whether whole or part-task training are needed, and the degree to which already trained skills can be transferred, are relevant in answering this question.

2.2.1. Amount of Initial Training

A consistent finding across many skill retention studies has been that the relative amount of initial training one receives is a strong predictor of level and duration of skill retention. Overtraining on one task may be considered insufficient training on a similar, analogous task. Also, overtraining can be very expensive if it requires significant high-fidelity simulator time. Due to the importance of this issue, reference is made to the figures presented earlier.

Using only a 15-step procedural task, Ammons et al. (1958, Experiment 1) trained subjects for either 5 or 30 training trials. Initial completion time for the two subject groups at the end of training was 0.4 and 0.2 minutes for the 5 and 30 trial groups, respectively. After a 2 year no-practice retention interval, the 5 trial training group performed the task in 1.3 minutes (a 3-fold increase) and the 30 trial group performance rose to 0.5 minute (2.5-fold increase). Proportionally fewer trials were required for the 5 trial subjects to regain their initial performance than were required by the 30 trial subjects. This training difference is plotted in Figure 1(B). In their tracking task (Ammons et al., 1958, Experiment 2), subjects were trained in aircraft control (using an airplane model) for a period of either 1 or 8 hours. Results from this study, plotted in Figure 2(A), showed that skills increased somewhat over retention intervals of up to 2 years. This increment was approximately equal for both training durations, but the 8 hour group maintained about a 2%-10% superiority over the 1 hour group performance throughout all retention intervals. While the superiority of longer training remained clear in this task, the reason for the performance increment did not.

Mengelkoch et al. (1960, 1971) trained inexperienced subjects for either 5 or 10 daily 50-minute sessions, in an aircraft flight simulator. As shown in Figure 1(C), the two groups had approximately the same skill degradation, after a 4 month retention interval, on a procedural task (losses for the 5 and 10 trial groups were, respectively, 20% and 16% of training levels). The effect of greater training was in achieving a nearly 20% increase in initial training level on the procedural task. The flight control or tracking portions of Mengelkoch's study only showed significant skill degradation, from both training groups, for the airspeed error parameter. The 5 and 10 trial training groups showed altitude error increases of about 10 feet over the 4 month interval (see Figure 3(A)), or about 20%-30% increase from initially trained levels. This loss was significant for the 5 trial group, but not the 10 trial group. Like performance of the procedural task, the primary difference between training duration groups was in the performance level at the end of training, rather than the relatively long skill retention. The fact that the 5 trial group retained their skills to the same magnitude as the longer trained group is meaningful.

Naylor et al. (1962, 1968) trained subjects on a dual tracking and procedural task for either 2 or 3 weeks of daily sessions. The longer training produced relatively superior performance at the end of both 1 and 4 week retention intervals when compared with end of training scores, but only in omissive errors. The comissive errors here did not significantly differ. Figure 1(F) shows the omissive errors from these subjects. Naylor et al. (1963) used the same task and trained subjects for either 5 or 10 daily sessions (one or two weeks). Only comissive errors here

differed as a function of training duration. The tracking performance from these dual-task studies showed similar trends. Three weeks of training produced less integrated tracking error and greater skill retention than two weeks of training. Integrated error was also significantly lower for 2 weeks of training, compared with 1 week of training. The 1 week trained group did, however, display increases in skill during the retention interval, not shown by the 2 week group.

Using a simple abstract line tracking task, Hammerton (1963) varied desired initial training criterions, as opposed to varying initial training durations. A 5% criterion group required 3 successive daily elapsed target acquisition times that did not differ at the 5% level of significance. Likewise, those in the 1% criterion group had 3 successive scores not differing at the 1% level. Retention of tracking skill after 6 months is shown in Figure 2(C). While two groups did not differ in mean time after training, the 5% group required more than 10 additional seconds than the 1% group after 6 months. This difference was both statistically and practically significant. Even the 1% group exhibited significant skill degradation, in spite of their extensive training. In this study, the additional training to achieve the 1% criterion was 9 to 17 days beyond the 8 to 22 days required for the 5% criterion. This degree of overlearning significantly decreased, but did not entirely alleviate, skill degradation.

Trumbo et al. (1965a) presented a similar line tracking task to 250 subjects. Half were trained for 50 trials and half for 100, over a 3 day interval. As shown in Figure 2(D), both training groups showed significant retention losses (increased tracking error) over a 5 month interval. The task organization was a stronger predictor of retention loss than was the absolute training duration. The high training level group did exhibit less skill loss than the low training group at all tested retention intervals. A subsequent analysis of separate skill retention components, from only the 100 training trial group, demonstrated that the best tracking subjects retained temporal accuracy (as measured by lead or lag time) better than spatial accuracy (as measured by percentage of over or undershoot errors). Thus, temporal as opposed to spatial training may be more important in retaining tracking skills over a long duration.

In a complex simulation of an Apollo mission, Youngling et al. (1968) trained their subjects for either 60 or 120 days. The overall skill retention, measured by time on target, was twice as great for the 120 trial group (5.5 seconds) than the 60 trial group (2.4 seconds on target).

Hagman (1983) summarized several skill retention studies performed in military contexts. Hagman (1980) varied the number of times Army personnel repeated a procedural electrical alternator output test during training. Increased task repetition, from 1 to 4 times, reduced performance time and errors by approximately constant amounts during training and after a two week retention. Increasing repetitions linearly increased performance until the 4 repetition duration training. Schendel and Hagman (1980) trained Army groups to either one correct performance or two correct performances in the disassembly and assembly of an M60 machine gun. After an 8 week retention interval, the greater trained group committed fewer errors than the lesser trained group. Goldberg et al. (1981) trained Army personnel to either 1 or 3 successive correct performances of boresighting and zeroing the main gun of an M60A1 tank. Again, higher performance after a 5 week retention was achieved by the more highly trained personnel. Schendel and Hagman also varied the time at which extra training was actually given. One group of subjects received extra task repetitions during the intitial training, while a second group received theirs half-way during the retention interval, at 4 weeks. They found no

significant difference between the two modes, implying that it is more cost effective to supply all training at one time.

Most of the above studies have been in agreement in that skill retention is a function of training duration. Many questions still remain, such as whether training duration is more or less important than task organization or the retention interval in determining the magnitude of skill retention.

2.2.2. Training Distribution

The way in which a given amount of training is distributed over a time interval is also predictive of skill retention success. Fleishman and Parker (1962) manipulated the method of retraining following a no-practice retention interval. A massed practice group received 4 practice sessions within a 2 hour period, while a distributed practice group received the same level of training, spread across 4 subsequent days. The distributed training group outperformed the massed practice group by the end of retraining, but both groups performed equally well after another 1 week retention interval. Thus, the distributed practice may have had its effect on temporary performance factors. Hagman and Rose (1983) reported that insertion of time between repetitions of a task increases skill retention, but the problems associated with the disruptive training may overshadow their benefits in actual tasks. Hagman (1980) compared massed versus spaced training for Army electrical alternator testers and repairers. The massed training group took 51% longer and made 40% more errors than the spaced group. Schendel and Hagman (1980) either gave task repetitions as part of initial training or one month later, and found no difference in ability after a two month retention interval. Spacing of task trials and/or sessions may be helpful, but there is some question as to whether its effectiveness varies with task proficiency level (Hagman and Rose, 1983). A model is needed here, and must consider the initial level of post-training skill proficiency, which also determines the required frequency of task repetitions.

2.2.3. Transfer of Training

Training transfer refers to the ability of a trained skill to generalize to a new setting. From cost considerations, positive skill transfer means that performance on a task can utilize already trained skills, saving time and money. Also, highly generalizable skills can easily be used in new settings or situations, for which no training has been constructed. The term validity refers to the degree to which training readies one for performance on a task, and is a measure of training transfer.

Briggs and Wiener (1966) trained subjects to perform an abstract two dimensional tracking task, and transferred this training to an easier task requiring the setting of a control knob. High fidelity training (achieved through proprioceptive control feedback) was only required when the transfer task reqiuired a high level of time sharing, by forcing constant positioning. Thus, when proprioceptive cues and high levels of time sharing are required in a task, the training program should be of high fidelity.

Reid (1975) assessed training transfer from a formation flight simulator to actual formation flying. Untrained, formation simulator trained, and aircraft trained pilots were compared in actual flight formation flying. Evidence of positive simulator skill transfer was obtained, as these

pilots did not fly significantly differently from conventionally trained pilots. The simulator provided the same degree of training as the flight sorties, indicating a high level of skill transfer. Carter and Trollip (1980) showed that training transfer between skills may be compared by plotting iso-transfer curves between pairs of skills and noting maximum transfer pairings. An operations research technique, the Lagrange Multiplier, was useful for determining costs and benefits of training.

Validity of training can also be evaluated by the method proposed by Goldstein (1978), who used a four level approach to evaluation: (1) Training validity, determined by trainee performance relative to standard training criteria, (2) Performance validity, measured by transfer of job performance, using criteria from the actual job, (3) Intra-organizational validity, measured by the performance of a group of new trainees based on the performmance of a previous group, and (4) Inter-organizational validity, measured by the degree to which a training program validated in one organization can be used in another organization. All of these levels must be evaluated to determine the effectiveness or validity of a given training program. Moving from the first level to the fourth, an increasing number of variables influence the success of training. Also, the necessary level of complexity in a training needs analysis must depend on the final goal of training. If one's goals do not reach beyond the second level, for instance, there is no need to consider levels 3 or 4. Such a structural assessment of validity is required to transform training needs assessment from art to engineering. The validity of training apparatus, according to Crawford and Crawford (1978), lies more in the manner in which it is used, rather than in the degree of its similarity to actual equipment. These investigators substituted conventional hands-on practice for part-task computer-based training on the use of an integrated control panel in an anti-submarine airplane. Control subjects performed on a high fidelity simulation of the control panel, while experimental subjects were trained using a graphic simulation on a touch screen display. The experimental subjects completed more tests, in less time than the control subjects. The computer-based training was found to provide at least as good skill acquisition, in less time and at lesser cost, than the full simulator training. A cost analysis indicated a substantial two-thirds cost savings over the conventional training method, much of which was due to a smaller number of instructor man-hours.

Adams (1979) contrasted the shortcomings of two methods of rating flight simulators for aircrew training. A transfer of training study measures the relationship between achieved task competence and proficiency on the flight simulator, while the rating method requires an engineering and experienced pilot assessment of hardware and flight similarity between the simulator and actual aircraft. Adams reviewed many studies with the thesis that both techniques are flawed. This does not mean that simulators are not useful, though. Humans require the perceptual-learning, stimulus-response learning, and feedback provided by simulator sessions. In addition, simulators successfully motivate trainees better than lower fidelity learning environments. Because simulators are based on these well-founded principles, simulators need not be evaluated for their effectiveness; this may be taken on faith (Adams, 1979).

As part of his procedural control setting training study, Johnson (1981) measured skill transfer by manipulating the sequential steps of the original task. In two experiments, low-fidelity paper-and-pencil training transferred very well to the new operational task. Although the two tasks likely utilized similar skills, this was further evidence of the utility of analogous tasks for training purposes. Validity determination and training needs assessment are still very much

debated topics; this section has only briefly introduced these issues.

2.2.4. Training Fidelity

A large and detailed lieterature has developed, concerning required fidelity in task training. At issue here is whether full fidelity is required to achieve adequate and retained skill performance. Considerations such as whether open or closed-loop simulation control is required, whether simulators must necessarily move, or whether adequate whole task performance can be achieved from part-task training have been addressed. In this paper, fidelity refers to the degree to which a training device can mimic an actual task of interest, such as flying an aircraft.

Naylor et al. (1963) manipulated the type of rehearsal subjects received in a dual tracking and procedural task. The procedural task consisted of 9 pairs of lights to which responses had to be made, and the tracking task was a three dimensional meter nulling task, in which roll, pitch, and yaw were simulated. In whole-task training, subjects practiced with both tasks simultaneously, as required for the measured performance task. Part-task training required separate practice on each task. Retention differences between the training conditions were significant on tracking performance, with the part-task rehearsal group displaying inferior performance. Whole-task rehearsal was also superior for procedural task performance, but this superiority lessened over the retention testing. Whole-task rehearsal was superior with a small amount of training (up to 5 days), but after 10 days of training, the two types of rehearsal were not signifianctly different. Naylor and Briggs (1963) manipulated rehearsal conditions on this procedural switch setting task. Whole-task rehearsal consisted of repeating the original task half-way through the 2 month retention interval. Part-task rehearsal conditions consisted of either (1) spatial rehearsal, with stimulus events occurring at equal temporal intervals, or (2) temporal rehearsal, with stimulus events occurring at varying times as in the original task, and stimuli appearing in a regular spatial order. The whole-task group produced far fewer omissive procedural errors than the part-task groups upon initial retention testing. The whole-task and part, temporal-task rehearsal were superior to spatial-task rehearsal when considering comissive errors. Thus, whole-task rehearsal here was best, closely followed by part-task temporal rehearsal in upholding skill retention over a 1 month retention interval.

Fleishman (1965) presented a multidimensional tracking task to inexperienced Air Force trainees, with the objective of predicting whole-task performance from various combinations of part-task training. The performance measurement device contained display dials for heading, altitude, bank, and airspeed, which all had to be simultaneously centered. Subjects were first proficiency trained on one dial, then two dials, then the entire task. Correlations between one-dial, part-task performance and whole-task performance ranged from .46 to .54 across the subjects. Between two-dial, part-task performance and whole-task performance, the range was .63 to .70. Multiple component practice was a better predictor of whole-task performance than single task performance in this multidimensional task. In addition, the multiple component performance was at least as predictive as linear combinations of the single task performances. The greatest correlations (.74) between part-task and whole task performance were found with linear combinations of two, two-dial practice. In this work, the actual task components that were used was less important than the fact that simultaneous practice had occurred. All predictive tasks here were part-task practice, but this investigation suggested that a continuum exists in training

effectiveness, between various integrative or combinatory levels of sub-task performances.

In his space vehicle approach and landing simulator, Sitterley (1974) varied the fidelity of pilot retraining methods, following a 4 month no-practice interval. The number of visual cues present in the training session strongly predicted the level of performance achieved, and the level of visual cueing was independent of the fidelity of the simulator session. Static photographic training was superior to open-loop, dynamic training. All cues, however, present in the static pictorial method were present in the dynamic display training, so the total number of cues was not soley predictive of training effectiveness. As stated by Sitterley, the most important element in these training alternatives was the presentation of efficient cues which assisted pilots in recalling their basic flight experiences. Thus, open-loop, static training methods may actually be superior to more costly methods, given careful training program design. Trollip (1979) compared a computer-based with a simulator-based training program for aircraft flight control. Control subjects were trained in a flight simulator, while experimental subjects were trained on a plasma touch screen terminal with an attached hand controller. The computer-based training produced significantly fewer critical errors and better flight control than the simulator. This trend was identical in both no wind and crosswind flight patterns. When generalizaing flight control to a new procedure, no difference was found between the two methods of training. The computer-trained subjects performed better, learned quicker, and made fewer mistakes than their control counterparts. It allowed students to develop better mental images of the ideal flight characteristics. Johnson (1981) also found that training requiring large use of mental imagery cues can produce the highest level of skill retention. Even in high complexity flight simulation and control environments, the highest level of fidelity is not required. Sitterley and Berge (1972) and Sitterley (1974) concluded that static rehearsal or training may be superior to the dynamic, higher fidelity rehearsal because of the artificially increased importance of visual cues.

One variable significantly influencing fidelity in aircraft simulations is flight motion. This has been a controversial topic over the past decade, with many insisting that motion cues are unnecessary for general aviator training. Caro (1979) discussed this issue with reference to two different motion cues, maneuver motion and disturbance motion. The former motion cue refers to those motion changes initiated by the pilot, whereas the latter refers to those cues initiated outside the immediate control loop, such as turbulance or engine effects. While maneuver motion moves the aircraft platform, it does not cause important alerting cues provided by the disturbance cues, which lead to quicker and more accurate pilot control of the simulator. No motion, on the other hand, is required if the simulated aircraft is easy to control and relatively stable (Caro, 1979). Thus, required fidelity here was based on a logical analysis of task training requirements.

In the monitoring and control of a procedural industrial operation, Johnson (1981) utilitzed three different training strategies: (1) conventional, full-fidelity practice on the actual task, (2) medium-fidelity reproduction study of photographs, where the subject was allowed to draw on the photos, indicating his procedural responses, and (3) low-fidelity, blind practice, where the subject was allowed to study, but not write on photos of the control equipment. Although the conventional strategy provided the quickest learning time (the blind practice required 1.5 times as long to reach criterion), the conventional and reproduction training did not produce different control setting errors after a 3 month retention interval. This illustrated that the highest fidelity training is not required in procedural tasks. Johnson and Rouse (1982) also found that low and medium fidelity training in aircraft power plant troubleshooting is very competitive to high

fidelity simulation. The highest-fidelity method in this study included training on the actual task, medium fidelity required a special power plant simulation, while the lowest fidelity condition utilized videotaped lectures and live quizzes. Video training produced the greatest performance with all simulated failures, and the original task and simulation were similar in their effectiveness. From a cost consideration, the low and moderate fidelity devices provided sufficient problem solving experience to effectively compete with the conventional training methods.

According to guidelines posed by Cream et al. (1978), the specification of required training fidelity appears to be art, rather than engineering. They stated that (1) essential and nonessential aspects of controls and displays must be differentiated, in that many of these elements are not required for proper training. (2) The choice of fidelity is more complicated when dealing with displays rather than simple indicators. Also, no rigorous decision-making procedures have been developed in the area of cost/benefit fidelity analyses. Though experiential-based fidelity definintion has been used for many years, no useful guidelines exist for the development of training for new tasks.

Perhaps the degree of required fidelity is a function of how little is understood of the processes required to carry out a given task. Seemingly complex tasks may only be combinations of a limited number of combined operations. On these, perhaps part-task training would be sufficient, if the actual components could be identified. Full-fidelity, whole-task training would then only be required in very complex tasks.

2.2.5. Adaptive Training

Both ground-based and on-orbit training systems should be adaptive to trainee performance, for maximum efficiency. This research area has recently shown substantial growth, as a result of the development of specific adaptive systems. Machine-controlled adaptive training simply automates a skilled instructor, by modifying the training stimuli as a function of trainee performance. Training efficiency is maximized, because effective learning only takes place when training is at an appropriate level of difficulty (Kelly, 1969). Adaptive learning curves typically show a linear relationship between ability and time, as opposed to conventional training curves.

The marker variable for training adaptation may vary, depending on the nature of a task. Johnson and Haygood (1984) utilized performance on a secondary light recognition task to adapt the difficulty of a primary tracking task. Williges and Williges (1978) concluded that the most effective adaptive parameter should be a multivariate combination of several performance skills. Matheny (1969) argued that the time lag between a system response and an operator's subsequent performance should serve as the adaptive parameter in general man-machine systems. While many parameters have been used, they have all served the function of varying the difficulty of a primary task.

2.3. RETENTION INTERVAL FACTORS

This section discusses those skill degradation issues directly related to the retention interval, between the initial training and actual performance. Two factors here have important implications for skill retention. (1) The duration of the retention interval has been extensively

studied, using intervals from a few minutes to well over two years. (2) The nature of activities performed during the retention interval influences skill degradation, just as practice of any task should do. Most investigators in this area have concluded that learned skills regularly degrade with increasing no-practice retention intervals. Bahrick (1964) has, however, disputed this concept, claiming that retention curves based on anticipation, recognition, or free recall reflect changes in one's sensitivity from session to session. This complaint, however, was only drawn against measures with only right/wrong responses. Many procedural studies, using other continuous measures such as completion time, have indeed found evidence for regular skill degradation over time.

2.3.1. Length of Interval

Retention intervals are an important consideration when designing refresher training on lengthy crew missions. The many skill retention/ degradation studies have all shown large performance decrements upon post-retention testing, when no retention practice is allowed. Estimates of the percentage skill loss at various time intervals allow an empirically determined estimate for the frequency of refresher training, given that retraining is required when performance falls below a set criterion. Longer intervals are generally accompanied by greater loss in skills, but this is very task specific. Ideally, this review should provide an overall skill retention function, mapping percentage of skills retained versus retention duration. However, reality dictates that between-study variations make such generalizations and models very hard to achieve. The most important question that can be answered here is whether a constant degradation across many tasks is found, with all other factors being equal. A cursory analysis of the results in Figures 1 through 3 indicates that skills degrade with time when not subjected to interim practice, and that the level of degradation reaches as asymptote in some studies.

Procedural Tasks. With few exceptions, procedural performance is marked by consistently increasing decrements with progressively longer retention intervals. Neumann and Ammons (1957) found that a one year, 90% loss in post-training performance was about the same as initial performance at the start of training, but proficiency was quickly regained upon retraining (Figure 1A). On a task of nearly equal complexity, Ammons et al. (1958, Expt. 1) found a 2 to 3-fold increase in task completion time after a one year interval, which did not appreciably increase after 2 years of retention (Figure 1B). The magnitude of relative skill degradation was the same here, regardless of the original number of training trials. Mengelkoch (1960; 1971) also found that relative mognitude of skill loss was independent of the amount of training (Figure 1C), where subjects showed a 20% decrease in correct procedures after a 4 month retention. In an extremely complicated 169 hour mission simulation, Cotterman and Wood (1967) found relatively small degradation over a 3 month retention when only a single parameter was considered (Figure 1D). The probability of successful performance over the interval fell by about .03. However, when all parameters in all phases of the simulation were considered, the probabilities dropped significantly over the interval; initially at an average of about 0.6, it fell to about 0.4 after the retention interval, suggesting that a failure was highly likely in some mission phases. This study was flawed, however, due to uncontrolled retention interval activities and small sample sizes (Gardlin and Sitterley, 1972). The performance of complex control and emergency procedures clearly degrade in required procedural time after 6 months, and Sitterley et al. (1972) noted a 4.5 fold increase after 4 months of retention (Figure 1E). Johnson (1981) measured the time required to set controls in an 87-step procedural task, and found a mean time

of 8 minutes after training had increased by 50% to 12.8 monutes after about 2.5 months. As gross estimates of the magnitude of skill degradation in procedural tasks, 20% to 50% degradation in 3 to 6 months, and 50% to 100% or more in more than 6 months may be made, based on the above data.

Investigations over shorter retention intervals, up to 1 month, have not found consistent skill degradation patterns. For example, using their switch setting task, Naylor and Briggs (1963) found a 20% decrease in omissive errors, but a 233% increase in comissive errors after 1 month (Figure 1F). Likewise, using the same task paired with a tracking task, Naylor et al. (1968) found the only retention degradation in comissive errors from the subject group with lesser training and low task organization (Figure 1F). Thus, skill retention of less than one month is harder to predict than longer durations, and may be dependent on many other task factors.

Simple Tracking Tasks. Performance on tracking tasks have not as a rule shown the predicatble and regular retention decrements shown by procedural task performance. In controlling the flight characteristics of a model airplane, Ammons et al. (1958, Expt. 2) found only a small 5% decrease in tracking time on target, between 1 and 24 months (Figure 2A). These slight skill decrements followed slight but significant increments between the end of training and 1 month retention. Beyond the absolute performance difference at training, the duration of training did not alter the relative decay rate of skills. Hammerton (1963) used a measure of target acquisition time in a tracking task (Figure 2C). By varying the allowable amount of session-to-session variability at the end of training, differential skill degradation was observed at a retention interval of six months. The looser criterion group showed a 3-fold increase in target acquisition time, whereas the tighter criterion showed only about a 1-fold increase. Thus, in tracking, regularity of performance as well as absolute magnitude appears to predict the degree of skill degradation. Over a short, 1 month retention, Naylor et al. (1962; 1968) demonstrated statistically significant skill loss at 2 levels of training duration and two levels of task organization (Figure 2E). Relative losses averaged about 16% at one week, and 44% at one month.

Four studies cited here used integrated tracking error as a dependent measure. Fleishman and Parker (1962) had two groups of subjects perform tracking tasks. A group receiving no formal training showed no performance decrement at up to 14 months of retention (Figure 2B). A second group who received formal training on the task showed a 1-fold increase in error after one year, but then showed a 4-fold increase after 2 years. Trumbo et al's (1965a) subjects showed virtually no performance decrement with intervals up to 5 months, when the task was unpredictable, with random targets located on every trial (Figure 2D). However, when the target position was more predictable, post-training tracking error was about 50% less than in the predictable condition. Retention intervals of 1 and 5 months produced large decrements in performance, upwards of 50%-60% from training levels. Equal degradation rates were found for both 50 and 100 training trial conditions, with the latter condition always producing better performance. Trumbo et al. (1965b) also demonstrated a 24% skill loss after a 1 month retention interval (Figure 2D). Swink et al. (1967) also manipulated task predictability and training duration in a tracking task (Figure 2F). The retention interval in this study, however, was unrelated to tracking error, as tracking ability did not degrade over a 5 month no-practice interval. Roehrig (1964) had several subjects stand on a small balancing platform, and measured the time duration that they could balance to within ± 1.5° of horizontal. After a 50 week hiatus from the task, all subjects demonstrated performance at least as great as shown at the end of

training. Much like the well-known fact that "one never forgets how to ride a bicycle," this task, once trained, seemed to trigger the same skill retention. Perhaps there is much ability transfer from balancing in ordinary walking (once the body has been trained to use the muscle groups required by the task), and subjects were unknowingly practicing the task. This conversly suggests that we can forget how to ride a bicycle, if balancing is not normally practiced, as in bed-ridden or spacefaring individuals. Research needs to first be conducted to determine which tasks are dependent on balancing practice in a gravity environment, as was implied by an earlier report (National Academy of Sciences, 1972, p. 245). A taxonomy of tasks, organized by gravity dependency, should be developed.

Simple tracking performance, not requiring a large number of simultaneous decisions and elements of concious cognitive control, does not appear to degrade as regularly and predictably as procedural skill performance. While some studies did find large decrements after a few months (e.g., Fleishman and Parker, 1962; Hammerton, 1963), others have found no evidence of skill degradation (e.g., Ammons et al., 1958; Swink et al., 1967). Clearly, in simple tracking, other factors are important in determining the degree of retention loss. From those studies cited here, those factors must include duration of initial training and task organization or predictability.

Complex Tracking Tasks. In those few studies using tracking tasks in higher complexity flight control contexts, performance on at least one parameter has shown strong interval-related degradation. Mengelkoch (1971) found significant increases in altitude error after a 4 month retention interval (Figure 3A). The skill degradation rate was equal between the 5 and 10 training trial groups, but the 10 trial group consistently made about 20% less error than the 5 trial group. Sitterley and colleagues also used altitude error, among many other parameters, in their space vehicle simulations. Sitterley and Berge (1972) measured a 2-fold increase in error over a 6 month interval, whereas Sitterley (1974) found a 5-fold increase over a 4 month retention interval (Figure 3B). In an alternate parameter, Sitterley and Berge measured a 55% increase in integrated pitch error after a no-practice retention of 6 months duration (Figure 3D). When measuring ability to null complex movements in the display within a simulator, Youngling et al. (1968) found a nearly linear relationship between the length of retention and performance loss (Figure 3C). Here, total tracking time on target decreased from approximately 40 seconds at training to about 33 seconds after 6 months, or a 20% loss. Percent flight skill degradation, a composite of many flight parameters, is perhaps the best overall measure of flight performance. Sitterley et al. (1972) noted a 400% decrease while Sitterley (1974) noted a 200% decrease in skills over a 4 month interval (Figure 3E). Clearly, flight skills are very sensitive to no-practice retention intervals, and may degrade by 4 or 5-fold over a few months.

2.3.2. Interpolated Activities

Practicing critical skills during the retention interval does aid retention performance. The relevant issues here are (1) for which types of tasks does practice aid, (2) what are the practice task transfer characteristics to the job performance task, and (3) are these dangers of negative task transfer; i.e., practice that can accelerate performance degradation.

Brown et al. (1963) required subjects to perform Naylor's switch setting task as well as a three dimensioal tracking task. Rehearsal on these tasks was manipulated on 4 days of a 15 day retention interval. For the tracking task, rehearsal greatly aided retention performance, but the

fidelity of the rehearsal did not alter this result. Performance decrements were effectively erased with task rehearsal. On the procedural task, rehearsal had influence on both commissive and omissive errors. For both types of tasks, sufficiently long original training attenuated the positive effects of rehearsal. When training was more limitied in scope, practice during the retention interval lead to large increases in skill retention. Naylor and Briggs (1963) tested variations in type of rehearsal, on retention performance of their switch setting task. The four rehearsal sessions occurred mid-way in their 25-day retention interval. One group received actual task rehearsal, one received no rehearsal training, and two groups received either part-task temporal or part-task spatial rehearsal. On omissive errors, the actual task group committed about half the errors of the other three rehearsal groups. On commissive errors, the actual task and temporal part-task groups were superior. Whole-task rehearsal was superior to part-task rehearsal conditions. Spatial rehearsal was barely any better than no rehearsal at all, but the time dimension may have been more difficult than the spatial dimension in this task. Trumbo et al. (1965b) compared verbal rehearsal with no rehearsal in a tracking task, over a one month retention interval. Part-task rehearsal required subjects to verbally repeat the tracking target location, referencing to its presented numerical location. On this task, no tracking mean performance retention difference was due to rehearsal, but a greater variability in tracking in the rehearsal group than non-rehearsal group was found.

Sitterley and colleagues investigated the type and distribution of rehearsal for their complex spacecraft simulation tasks. Sitterley and Berge (1972) presented both emergency procedural and flight control tasks to thier inexperienced subjects. After four months of inactivity, both task performances were greatly degraded, beyond the minimal proficiency level. As part of their experimental design, two subject groups recieved static rehearsal training during the retention interval period, where a session consisted of a review of the flight training manual, photographs of the cockpit environment, and a written evaluation test. The static rehearsal greatly countered performance degradation for the procedural task, at both 3 and 6 month intervals. The interim rehearsal aided performance as much as allowing dynamic warmup immediately prior to the start of retention testing. The continuous task, on the other hand, responded differently to rehearsal and warmup training. At a 6 month retention interval, static rehearsal was insufficient to maintain performance in all control skills; dynamic warmup was required to insure reliability. The regular rehearsal sessions were, however, adequate for skill maintenance over the 3 month retention interval. Thus, long retention intervals require both rehearsal and warmup for flight control, but only require static rehearsal for procedural tasks. Using an even more complex space vehicle approach and landing under both visual and instrument flight conditions, Sitterley et al. (1972) added a condition of dynamic rehearsal training, in addition to improving the static rehearsal training method. The improved static method utilized photographs of flight instruments and scenes at critical times, and allowed the subject to sit in the simulator cockpit for refamiliarization prior to testing. The dynamic rehearsal condition included the above static rehearsal, then the pilots were allowed to view three dynamic flights from the cockpit, in an open-loop fashion. The pilot still did not have direct interaction as he would have during warmup practice. Results showed that, like Sitterley and Berge (1972), static rehearsal improved skill retention, but required dynamic warmup practice for adequate proficiency. The dynamic rehearsal prevented skill degradation for all procedural and flight control tasks, with the visual flight control portions receiving the greatest training benefit. The static method was only slightly worse than the dynamic method in retention of flight control or continuous skills.

From a cost/benefit viewpoint, Sitterley (1974) claimed that the static rehearsal method had the greatest development potential. To test an advanced static rehearsal version, Sitterley presented more pictorial information along both normal and sub-nominal flight paths, and enhanced pilot involvement to reinforce critical perceptual cues in the visual environment. The static rehearsal was presented in a booklet format for self-study by the pilot. All retention testing was preceded by a 40 minute slide show of real time cockpit views of the approach and landing. After a 4 month retention interval, the advanced static retraining countered all skill degredation, more so than even the dynamic rehearsal of the previous study. Sitterley suggested that the carefully structured visual cues at critical moments were sufficient to key appropriate pilot responses.

The above studies have nearly all confirmed the utility of retention interval practice in countering both continuous and procedural skill degradation. When training is of insufficient duration, rehearsal methods can be substituted to some degree. The rehearsal training should be carefully designed to provide minimal cues required to successfully perform the task of interest. Experiments by Sitterley have demonstrated that rehearsal for complex flight control does not have to be closed loop and high fidelity. So long as the important visual cues have been provided, open-loop, pictorial reviews may adequately be substituted for the real task. Of those studies reported here, none have concluded that rehearsal degrades retention when compared with no rehearsal. However, none have systematically varied rehearsal tasks so as to provide negative skill transfer.

3. SPACE MISSION TRAINING

It is frustrating to study empirical research on task training factors, then consider the techniques that are actually used to define training requirements. Cream et al. (1978) outlined "systematic methods" usually used to specify training objectives in a specific task. First, the behavioral skills and knowledge required of graduated trainees are identified. Next, these are matched against the actual ability of new students. The identified differences then define training requirements of a program. As recognized by Cream et al., a lack of task analysis data for defining training requirements exists, especially with new systems. They recommended seeking out analogous tasks, again avoiding the issue of task analysis. Such experiential-based development can be a costly error in new systems development, where many competing task factors can eclipse unforseen interactions. This report defends the need for a quantitative model of training requirements as a function of task factors. Of course, much research will be required to specify this model.

This section presents a brief outline of operational space mission training at NASA, for the purpose of establishing a foundation from which to define Mars mission training requirments.

3.1. CURRENT MISSION TRAINING

Training programs for Space Shuttle missions proceed from part-systems teaching and practice to more complex, fully integrated simulations. A typical astronaut training program currently requires about 5 years from start to flight. Training starts with stand-alone equipment, then proceeds to joint, integrated mission simulation. Initially, workbooks and self-paced computer aided trainers are used to gain knowledge and proficiency on specific systems. Single and multiple

part-task trainers are then used to gain required psychomotor proficiency. Examples of these trainers are specific shuttle control panels and the RMS. An underwater weightless training facility, and airborne parabolic flights may be used for specific procedure training. Shuttle systems simultators are now used for many tasks. The multiple-task shuttle simulator may be tied with one or more flight centers in partial-mission simulations. The full mission may then be simulated by tying in all payload customers and control centers. For very complicated maneuvers, this joint simulation may even be repeated on-orbit just prior to the actual performance.

Very little active astronaut performance measurement is currently conducted, once selected (Akins, 1979; Nicogossian, 1984; Christensen and Talbot, 1985). Throughout the literature, a prerequisite for the evaluation and developpment of training procedures is unbiased performance data (see Goldstein, 1978; Cream et al., 1978; Swezey, 1979). As an illustration of this, consider a part-mission simulator session. Trainers prepare scripts of system failures that occur at regular intervals, every few minutes. The task of the trainees is to make educated diagnoses and decisions, while controlling the space vehicle. After completing each simulation run, the training scripts are reviewed with the trainees, pointing out mistakes that were made. A new run then begins, with the hope that lessons have been learned. While trainees clearly learn from this training scenario, limitations in vehicle design or human capabilities are not collected. A series of failures made by all trainees would not be noted; such failures are valuable data to use in the redesign of systems. A separate performance monitoring system, invisible to trainers and trainees, would be useful here.

Soviet cosmonauts also utilize simulators and part-task trainers, but their training philosophy differs in a basic way from U.S. philosophy. Rather than rely on basic documentation in training programs, they listen to a lecture from a specialist several times, taking notes (Lenorovitz, 1982). The Soviet program also places more emphasis on psychological status than the U.S. program, with tests given at training to assure psychological compatibility with crew members, and regular psychological monitoring during and after flight (Borrowman, 1982; Bluth, 1982; Oberg and Oberg, 1986). Future U.S. missions must concentrate more on psychological status of crew members during training (Conners et al., 1985; Collins, 1985).

3.2. SPACE STATION TRAINING

A recently published document (NASA, 1986) detailed training requirements for the near-operatioal space station, to be launched in the early 1990's. This section will summarize important aspects of this paper.

3.2.1. On-Orbit Versus Ground-Based Training

Specific criteria have been imposed to assign training to on-orbit or ground-based trainers. On-orbit training is preferred for complex psychomotor skills, or time-critical procedures, safety drills, and maintenance of group behavioral dynamics. This training is preferred for microgravity, low-cost, and low probability of occurrance tasks. On-orbit refresher training will also be carried out prior to unscheduled maintenance tasks. Ground-based training will be preferred for fundamental, safety-critical tasks, such as space station activation or medical procedures. Basic training in group dynamics and habitability will be carried out on the ground.

Ground-based training will, in general, be preferred for prerequisite skill acquisition and day-to-day space station operations. On-orbit training is a supplament to basic ground-trained skills, so few skills will only be trained on-orbit.

Initially, most training will be performed on the ground, but more and more training will be performed on-orbit as the space station program matures. Ground-based training will include (1) ingress and egress to and from space station, (2) activation and deactivation of space station, (3) systems training, with emphasis on understanding, (4) spacecraft docking and tethering, (5) RMS and robotics, (6) orbital management and communications, (7) habitability systems, (8) safety, emergency, medical and maintenance procedures, (9) integrated simulations, stressing team approach to problems. On-orbit training will include (1) spacecraft docking and tethering, (2) refresher training on RMS, (3) crew rescue EVA, (4) handling of fuels and other hazardous materials, (5) use of avionics equipment, (6) emergency and malfunction procedures, (7) manned systems refresher training. Ground-based training will initially be performed for all phases and activities. Eventually, primary training for some skills or activities will be shifted to on-orbit. This training will be available both to NASA astronauts and to contractors or customers of NASA.

3.2.2. Training Breadth

Training programs for the space station will be developed for the crew, ground-based flight controllers, and training instructors. The crew members will require more training at greater frequency for procedural skills than for psychomotor skills. Launch schedules will impose training duration limits. There are still questions as to the relative amount of self-paced training, amount of on-board training, and relative training differences between crew members. Flight controllers will initially take part in full integrated mission simulations, however later in the space station program, fewer formal simulations will be conducted. Eventually, no joint controller-crew simulation will take place, due to their inherent complexity and time consumption. Instructors must also be trained in procedures. Questions exist as to the number of required instructors per crew member, and the complexity of their simulation scripts.

The general direction for breadth of training is one of initial full scale, integrated simulations involving all parties tapering to later separate simulations of mission components. This change will be required to shorten the training time of space station crews, and to decrease the cost of rotating crews. Some amount of procedural or psychomotor skill practice, such as one-half hour per day, will be mandatory on-orbit.

3.2.3. Training Technologies and Facilities

Space station training will make extensive use of computer aided and adaptive instruction. Computer aided instruction systems permit a consistent presentation of material in a given sequential order. Intelligent or adaptive computer aided instruction allows material to be resequenced or altered according to the needs of a trainee (Morgan and Erb, 1986). Both of these types of systems will be utilized in ground-based and on-orbit training. Intelligent systems will be designed to serve as a coach, rather than as a tutor or manager, in that advice is provided to the trainee trying to meet an educational objective.

These instructional programs will be implemented on interactive laser videodisk storage systems. The trainee will respond via keyboard and voice. The system will output via television monitors, wide angle visuals, helmet mounted video, and voice synthesis. Appropriate videodisks for every required repair or maintenance will be onboard; the capacity for transmitting the equivalent content of a videodisk directly into the training hardware may also be present. These systems offer many advantages over conventional computer-based trainers. They are potentially very small and portable, still or moving scenes are of higher fidelity than computer graphics, and it is cheaper to film sequences of movements for videodisk interpretation than developing reliable graphics via computer. Astronauts on EVA will have the capability of viewing procedures in a helmet mounted display as they are performed. These videodisk trainers can also contain other controller attachments to allow realistic practice with complex psychomotor skills.

Space station training will also be embedded within operational controls and displays. By monitoring performance while a trainee attempts to complete a given task, better involvement and motivation are achieved. The monitor will act as a coach, much like a master-apprentice scenario.

Ground-based NASA training facilites for the space station missions will include (1) manned test facility, for engineering design and testing, (2) mockup and integration laboratory, for RMS training, (3) weightless environment training facility, for crew training, (4) systems engineering simulator, for flight training, (5) space station training facility, for high-fidelity crew training, (6) shuttle rendezvous simulator, and (7) integrated EVA simulator. As in the shuttle training program, training will start with part-task, single-system trainers, and end with full-mission, multiple system simulation. Much of the shuttle training facilities will be utilized for space station training.

3.4. MARS MISSION TRAINING

Until an empirically determined model of training requirements has been developed, all conceptual designs are merely "straw-man" estimates. However, based on previous sections of this paper, some recommendations may be made.

On-Board Versus Ground-Based Training. Whether training should be given on-board or on the ground should not be an issue on a Mars mission. The crew members must have the necessary resources to rehearse procedural lists and psychomotor skills whenever required. Ground-based training should consist of academic systems overviews and those skills required for complete systems understanding. It is imperitive that complete understanding be achieved prior to flight, as effective refresher training can only be assured with well organized tasks. Ground-based training might include other knowledge acquisition, beyond the immediate scope of the mission, to guard against unexpected events. As an example, this training might include psychological or social models of small groups. In general, ground-based training should be academic and broad, while on-board required training shoud be specific and skill oriented. Of course, pilots will require all training prior to the mission.

Scope of Training. Ground-based and on-board training and refresher programs must be designed to counteract the negative aspects of the space environment, as discussed in Section 1, in addition to maintaining skill and knowledge retention. Skills must be regularly refreshed, according to a

yet-to-be developed model of retention time. Using such a model, a computer program could list, on a daily basis, those skills or procedures that need refresher training. Ideally, a crew member's required refresher training should be determined automatically and adaptively. Periodic performance measurement on a testing battery could indicate level of retention and pinpoint areas for needed training. Training for crew autonomy and confinement will be harder to define, until more is known. Drills may be required to measure the cohesiveness of the crew. Lists of critical procedures must be regularly reviewed and trained, as should the daily workload level. As measured by a model, mental and physical workload must be constantly reviewed and reallocated among the crew members.

All crew members should be encouraged to develop expertise, while in-transit to Mars, in academic fields other than their own. The on-board teaching expertise clearly will exist. Establishing a formal instructional regimen will aid in maintaining cognitive abilites of both teachers and students. Healthy interaction between the crew members will also be maintained. The outcome of such concentrated training could even consist of additional academic degrees.

Periodic Drills. Emergency and disaster drills should be conducted, as called for by either ground control or by the on-board commander. Many controls could be placed in an alternate, embedded training mode for conducting these drills. Images of impending meteorites, etc., could even be projected onto displays or windows. Crew performance should be reviewed by the commander, and necessary refresher training conducted.

Recreation. Off-duty periods also present a good oppurtunity for procedural and operational skill maintenance. Video games, music, etc., all present unique practice oppurtunities for different skills.

Hardware. A small, portable, videodisk based computer system with voice input and output may serve as a generic trainer and recreation device. Such a device will allow practice of skills anywhere and anytime on a mission. Different videodisks could be loaded for different procedures, and others could be loaded for entertainment.

A Mars mission presents many challenges far beyond those that have already been approached. This oppurtunity should be seized for pushing the state-of-the-art in knowledge of human training and skill retention. This paper has stressed the development of empirical models, the only unbiased approach to defining training needs. As the research required to achieve these models will take many years, now is the time to start. A long duration space mission will require an understanding of psychological limitations in all mission phases. This report has stressed the need for modeling these limitations in light of training requirements, whether initial or refresher training. The proper, scientific method of training definition will require a model of skill retetention, as argued here.

4. REFERENCES

Adams, J.A., "On the Evaluation of Training Devices," Human Factors, 1979, 21(6): 711-720.

Akins, F.R., "Performance Considerations in Long-Term Space Flight," NASA Technical Report N81-18651, Contract No. NASA-CR-152384, 1979.

Alluisi, E.A., "Methodology in the Use of Synthetic Tasks to Assess Complex Performance," Human Factors, 1967, 9(4): 375-384.

Ammons, R.B., Farr, R.B., Bloch, E., Neumann, E., Day, M., Marion, R., and Ammons, C.H., "Full-Term Retention of Perceptual Motor Skills," Journal of Experimental Psychology, 1958, 55:318-328.

Bahrick, H.P, "Retention Curves: Facts or Artifacts?," Psychological Bulletin, 1964, 61(3): 188-194.

Bahrick, H.P., "The Ebb of Retention," Psychological Review, 1965, 72: 69-73.

Bluth, B.J., "Staying Sane in Space," Mechanical Engineering, 1982, 1: 24-29.

Borrowman, G.L., "Training for Space," Spaceflight, 1982, 24: 129-132.

Brown, I.D., "Dual Task Methods of Assessing Workload," Ergonomics, 1978, 21(3): 221-224.

Brown, D.R., Briggs, G.E., and Naylor, J.C., "The Retention of Discrete and Continuous Tasks as a Function of Interim Practice with Modified Task Requirements," Wright-Patterson AFB Technical Report AMRL-TDR-63-35, Ohio, 1963.

Briggs, G.E., and Wiener, E.L., "Influence of Time Sharing and Control Loading on Transfer of Training," Journal of Applied Psychology, 1966, 50(3): 201-203.

Caro, P.W., "The Relationship between Flight Simulator Motion and Training Requirements," Human Factors, 1979, 21(4): 493-501.

Carter, G., and Trollip, S., "A Constrained Maximization Extension to Incremental Transfer Effectiveness, or, How to Mix your Training Technologies," Human Factors, 1980, 22(2): 141-152.

Chiles, W.D., "Methodology in the Assessment of Complex Performance: Discussion and Conclusions," Human Factors, 1967, 9(4): 385-392.

Christensen, J.M., and Talbot, J.M., "Research Opportunities in Human Behavior and Performance," NASA Contractor Report 3886, Contract NASW-3924, 1985.

Collins, D.L., "Psychological Issues Relevant to Astronaut Selection for Long-Duration Spaceflight," Brooks AFB Technical Report AD-A154 051, Texas, 1985.

Conners, M.M., Harrison, A.A., and Akins, F.R., Living Aloft, Human Requirements for Extended Spaceflight, Washington, D.C.: NASA, 1985.

Cotterman, T.E., and Wood, M.E., "Retention of Simulated Lunar Landing Mission Skill: A Test of Pilot Reliability," USAF Technical Report AMRL-TDR-66-222, Wright-Patterson AFB, Ohio, 1967.

Crawford, A.M., and Crawford, K.S., "Simulation of Operational Equipment with a Computer-Based Instructional System: A Low Cost Training Technology," Human Factors, 1978, 20(2): 215-224.

Cream, B.W., Eggemeier, F.T., and Klein, G.A., "A Strategy for the Development of Training Devices," Human Factors, 1978, 20(2): 145-158.

Fleishman, E.A., "The Prediction of Total Task Performance from Prior Practice on Task Components," Human Factors, 1965, 7(2): 18-27.

Fleischman, E.A., and J.F. Parker, "Factors in the Retention and Relearning of Perceptual-Motor Skills (b)," Journal of Experimental Psychology, 1962, 64: 215-226.

Fraser, T.M., "Leisure and Recreation in Long Duration Space Missions," Human Factors, 1968, 10(5): 483-488.

Gardlin, G.R., and Sitterley, T.E., "Degradation of Learned Skills, A Review and Annotated Bibliography," Report No. D180-15080-1, The Boeing Co., (NASA acquisition No. N73-10152, contract No. NAS9-10962), Seattle, WA, 1972.

Garvey, W.D., "A Comparison of the Effects of Training and Secondary Tasks on Tracking Behavior," Journal of Applied Psychology, 1960, 44(6): 370-375.

Goldberg, S.L., Drillings, M., and Dressel, J.D., "Mastery Training: Effects on Skill Retention," U.S. Army Research Inst. Behav. Soc. Sci. Technical Report 513, 1981, Alexandria, VA.

Goldstein, I.L., "The Pursuit of Validity in the Evaluation of Training Programs," Human Factors, 1978, 20(2): 131-144.

Gopher, D., and North, R.A., "Manipulating the Conditions of Training in Time-Sharing Performance," Human Factors, 1977, 19(6): 583-593.

Grimsley, D.L., "Acquisition, Retention and Retraining: Effects of High and Low Fidelity in Training Devices," Human Resources Research Office, Technical Report 69-1, 1969.

Grodsky, M.A., and Lutman, C.C., "Pilot Reliability and Skill Retention for Space Flight Missions," Air University Review, 1964, 16: 22-32.

Hagman, J.D., "Effects of Training Task Repetition on Retention and Transfer of Maintenance Skill," U.S. Army Research Inst. Behav. Soc. Sci. Technical Report 1271, 1980, Alexandria, VA.

Hagman, J.D., and Rose, A.M., "Retention of Military Tasks: A Review," Human Factors, 1983, 25(2): 199-213.

Hammerton, M., "Retention of Learning in a Difficult Tracking Task," Journal of Experimental Psychology, 1963, 66:108-110.

Joels, K.M., The Mars One Crew Manual, New York: Ballantine Books, 1985.

Johnson, D.F., "Research Note: The Use of Secondary Tasks in Adaptive Training," Human Factors, 1984, 26(1): 105-108.

Johnson, S.L., "Effect of Training Device on Retention and Transfer of a Procedural Task," Human Factors, 1981, 23(3): 257-272.

Johnson, W.B., and Rouse, W.B., "Training Maintenance Techniques for Troubleshooting: Two Experiments with Computer Simulations," Human Factors, 1982, 24(3): 271-276.

Kelly, C.R., "What is Adaptive Training?," Human Factors, 1969, 11(6): 547-556.

Lenorovitz, J.M., "Soviet Cosmonauts Training at Star City Complex," Aviation Week & Space Technology, August, 1982, 44-46.

Lincoln, R.S., "Learning and Retraining a Rate of Movement with the Aid of Kinesthetic and Verbal Cues," Journal of Experimental Psychology, 1956, 51(3): 199-204.

Matheny, W.G., "The Effective Time Constant - A New Technique for Adaptive Training," Human Factors, 1969, 11(6): 557-560.

Mengelkoch, R.F, Adams, J.A., and Gainer, C.A., "The Forgetting of Instrument Flying Skills as a Functio of the Level of Initial Proficiency," NAVTRADEVCEN Technical Report 71-16-18, New York, 1960.

Mengelkoch, R.F, Adams, J.A., and Gainer, C.A., "The Forgetting of Instrument Flying Skills," Human Factors, 1971, 13(5): 397-405.

Moray, N., "Models and Measures of Mental Workload," in Moray, N. (Ed.), Mental Workload, Its Theory and Measurement, New York: Plenum Press, 1979.

Morgan. L.W., and Erb, D., "Survey of Current Training Technology," NASA Working Paper, Crew Training Division, Johnson Space Center, April, 1986.

NASA, Space Station Training Document, Crew Training Divison, Johnson Space Center, 1986.

National Academy of Sciences, Human Factors in Long-Duration Spaceflight, Washington, D.C.: National Academy of Sciences, 1972.

National Commission on Space, Pioneering the Space Frontier, Toronto: Bantam Books, 1986.

Naylor, J.C., and Briggs, G.E., "Long-Term Retention of Learned Skills: A Review of the Literature," USAF Technical Report ASD 61-390, Behavioral Sciences Laboratory, Wright-Patterson AFB, Ohio, 1961.

Naylor, J.C., and Briggs, G.E., "Effect of Rehearsal of Temporal and Spatial Aspects on the Long-Term Retention of a Procedural Skill," Journal of Applied Psychology, 1963, 47(2): 120-126.

Naylor, J.C., Briggs, G.E., and Buckout, "Long-term Skill Transfer and Feedback Conditions during Training and Rehearsal,"USAF Technical Report AMRL-TDR-63-136, Wright-Patterson AFB, Ohio, 1963.

Naylor, J.C., Briggs, G.E., Brown, E.R., and Reed, W.G., "The effect of Rehearsal on the Retention of a Time-Shared Task," USAF Technical Report, AMRL-TDR-63-33, Wright-Patterson AFB, Ohio, 1963.

Naylor, J.C., Briggs, G.E., and Reed, W.G., "Task Coherence, Training Time, and Retention Interval Effects on Skill Retention," Journal of Applied Psychology, 1968, 52: 386-393.

Neumann, E., and Ammons, R.B., "Acquisition of Long-Term Retention of a Simple, Serial, Perceptual-Motor Skill," Journal of Experimental Psychology, 1957, 53: 159-161.

Nicogossian, A.E., "Human Capabilities in Space," NASA Technical Memorandum 87360, Washington, D.C., 1984.

Noble, M., and Trumbo, D., "The Organization of Skilled Response," Organizational Behavior and Human Performance, 1967, 2: 1-25.

Oberg, J.E., Mission to Mars, Plans and Concepts for the First Manned Landing, New York: Meridian, 1982.

Oberg, J.E., and Oberg, A.R., Pioneering Space, Living on the Next Frontier, New York: McGraw-Hill Co., 1986.

Parker, J.F., and Fleishman, E.A., "Use of Analytical Information Concerning Task Requirements to Increase the Effectiveness of Skill Training," Journal of Applied Psychology, 1961, 45(5): 295-302.

Reid, G.B., "Training Transfer of a Formation Flight Trainer," Human Factors, 1975, 17(5): 470-476.

Roehrig, W.C., "Psychomotor Task with Perfect Recall After Fifty Weeks of No Practice," Perceptual and Motor Skills, 1964, 19: 547-550.

Savage, R.E., Williges, B.H., and Williges, R.C., "Empirical Prediction Models for Training-Group Assignment," Human Factors, 1982, 24(4): 417-426.

Schendel, J.D., and Hagman, J.D., "On Sustaining Procedural Skills over Prolonged Retention Intervals," U.S. Army Research Institute for the Behavioral and Social Sciences Technical Report 1234, 1980, Alexandria, VA.

Shields, J.L., Goldberg, S.L., and Dressel, J.D., "Retention of Basic Soldiering Skills," U.S. Army Research Institute for the Behavioral and Social Sciences Technical Report 1225, 1979, Alexandria, VA.

Sitterley, T.E., "Degradation of Learned Skills, Static Practice Effectiveness for Visual Approach and Landing Skill Retention," Report No. D180-17876-1, The Boeing Co., (NASA contract No. NAS9-13550), Seattle, WA, 1974.

Sitterley, T.E., and Berge, W.A., "Degradation of Learned Skills, Effectiveness of Practice Methods on Simulated Space Flight Skill Retention," Report No. D180-15081-1, The Boeing Co., (NASA acquisition No. N73-10159, contract No. NAS9-10962), Seattle, WA, 1972.

Sitterley, T.E., Zaitzeff, L.P., and Berge, W.A., "Degradation of Learned Skills, Effectiveness of Practice Methods on Visual Approach and Landing Skill Retention," Report No. D180-15082-1, The Boeing Co., (NASA acquisition No. N73-23086, contract No. NAS9-10962), Seattle, WA, 1972.

Swezey, R.W., "An Application of a Multi-Attribute Utilities Model to Training Analysis," Human Factors, 1979, 21(2): 183-189.

Swink, J., Trumbo, D., and Noble, M., "On the Length-Difficulty Relation in Skill Performance," Journal of Experimental Psychology, 1967, 74: 356-362.

Trollip, S.R., "The Evaluation of a Complex Computer-Based Flight Performance Trainer," Human Factors, 1979, 21(1): 47-54.

Trumbo, D., Noble, M., Cross, K., and Ulrich, L., "Task Predictability in the Organization, Acquisition and Retention of Tracking Skills," Journal of Experimental Psychology, 1965a, 70: 252-263.

Trumbo, D., Ulrich, L., and Noble, M., "Verbal Coding and Display Coding in the Acquisition and Retention of Tracking Skill," Journal of Applied Psychology, 1965b, 49: 368-375.

Trumbo, D., Noble, M., and Swink, J., "Secondary Task Interference in the Performance of Tracking Tasks," Journal of Experimental Psychology, 1967, 73: 232-240.

von Tiesenhausen, G., "An Approach Toward Function Allocation Between Humans and Machines in Space Station Activities," <u>NASA Technical Memorandum 82510</u>, Marshall Space Flight Center, Alabama, 1982.

Williges, R.C., and Williges, B.H., "Critical Variables in Adaptive Motor Skills Training," <u>Human Factors</u>, 1978, 20(2): 201-214.

Youngling, E.W., Sharpe, E.N., Ricketson, B.S., and McGee, D.W., "Crew Skill Retention for Space Missions up to 200 Days," <u>Report F7666, McDonnell-Douglas Astronautics Co.</u>, Eastern Div., 1968.

1986

NASA/ASEE SUMMER FACULTY FELLOWSHIP PROGRAM

Johnson Space Center

University of Houston

Effect of STS Space Suit on Astronaut Dominant Upper Limb

EVA Work Performance

Prepared by: Michael C. Greenisen

Academic Rank: Associate Professor

University and Department: University of Wisconsin-Milwaukee

                                          Human Kinetics

NASA/JSC

Directorate: Space and Life Sciences

Division: Man System

Branch: Design and Analysis

JSC Colleague: James Taylor

Date: December 17, 1986

Contract #:   NGT-44-005-803

                                          University of Houston

"EFFECT OF STS SPACE SUIT ON ASTRONAUT DOMINANT UPPER LIMB EVA WORK

PERFORMANCE", M. G. Greenisen, Ph.D., Anthropometric Lab, NASA, SP22,

Johnson Space Center, Houston, Texas 77058.

The purpose of this investigation was to evaluate STS Space Suited (3.7

PSID) and unsuited dominant upper limb performance in order to quantify

future EVA Astronaut skeletal muscle upper limb performance expecta-

tions. Testing was performed with subjects standing in EVA STS foot

restraints. Data was collected with a CYBEX Dynamometer enclosed in a

waterproof container. Control data was taken in one "g". During one

"G" testing weight of the Space Suit was relieved from the subject via

an overhead crane with a special connection to the PLSS of the suit.

Experimental data was acquired during simulated zero "g", accomplished

by neutral buoyancy in the Weightless Environment Training Facility.

Unsuited subjects became neutrally bouyant via SCUBA BC vests. Actual

zero "g" experimental data was collected during parabolic arc flights

on board NASA's modified KC-135 aircraft. During all test conditions

subjects performed five EVA work tasks requiring dominant upper limb

performance and ten individual joint articulation movements. Dynamome-

ter velocities for each tested movement were 0 degree/second, 30 or 60

degree/second and 120 or 180 degree/second, depending on the test, with

three repetitions per test. Performance was measured in foot pounds of

torque. Testing continues.

INTRODUCTION

This report represents an "in progress review" of an ongoing Space Biomechanics research project with NASA JSC, Directorate of Space and Life Sciences. The study was conceptualized and initiated during a NASA-ASEE Faculty Fellowship, Summer 1985. This research continued full time at NASA JSC, September 1985 - August 1986 under a university academic year sabbatical leave and a second NASA-ASEE Faculty Fellowship award. Since August 1986 the study has been supported by a Lockheed - NASA Contract, number NAS9-15800 dated 18 August 1986.

The basis of this report is extracted from documentation prepared for a Weightless Environment Training Facility (WETF) Test Readiness Review Board (TRRB), which was held on August 21, 1986. The TRRB approved the use of the Cybex Dynamometer for experimental procedures in the WETF. Essentially, the TRRB documentation is presented in total.

This documentation summarizes the engineering effort and experimental design evolution from June 1985 to date. The study will continue through the Summer 1987. A final report on the results of this study will be published with the Summer 1987, NASA-ASEE Faculty Fellowship final reports from NASA JSC.

RESEARCH CONCEPT

This investigation is evaluating STS space suited and unsuited dominant upper limb performance in order to quantify future EVA upper limb work performance expectations. Data is being generated by Mission Specialist Astronauts. Data is collected during one control and two experimental environments.

To standardize testing, all data is being collected with a Cybex dynamometer enclosed in a nitrogen charged waterproof container pressurized to 4 psid, mounted on a dedicated stand. The stand meets all test environment attachment specifications and also serves as the anchor point for astronaut STS foot restraints.

Control one-g data is collected in the Anthropometric Measurement Laboratory (AML). During one-g testing, weight of the space suit is relieved from the subject via an overhead crane specially connected to the PLSS of the suit.

Experimental data is acquired during simulated zero-g accomplished by submerging the subjects and Cybex in the Weightless Environment Training Facility (WETF). Suited subjects become neutrally bouyant (NB) through standard WETF ballast weighting system techniques. Unsuited subjects become NB via SCUBA BC vests.

Actual, zero-g experimental data is collected during parabolic arc flights on board NASA's modified KC-135 aircraft. A gravity meter measures actual negative "g" level attained. During all dedicated test conditions, subjects perform five EVA work tasks requiring whole dominant upper limb performance and ten individual joint articulation movements of the same limb. Dynamometer velocities vary depending on the test. However, 0 degree/second is utilized during all tests to determine levels of inertial forces which may be manually exceeded by STS suited astronauts.

Data generated by a prototype hand held force measuring unit from Scott Science and Technology is also being compared to Cybex results during selected protocols of this study. This hand held unit is not waterproofed and will not be part of the Cybex WETF testing at this time.

The efforts of this study, to date, have focused on the design engineering and fabrication of the required hardware, plus evolution of the research experimental design. Principal Investigator familiarization with KC-135 parabolic flights and NASA WETF SCUBA qualification has also occurred. Research pilot data has been collected with the instrumentation on four KC-125 zero-g flights. Two flights were with unsuited and two flights were with suited astronaut subjects. During these flights the instrumentation functioned without problems.

Collection of data began the first week of November from three STS suited astronauts which submerged in the WEFT. The Cybex dynamometer, enclosed in the waterproof container, performed without problems as the test article. After testing, the underwater container was unsealed and inspected. No evidence of water leakage into the container was found, thereby, documenting the success of the engineering design.

There is inadequate data, as of this date, to realistically present any findings. However, two areas of interest have surfaced based on the available data: 1) the range of motion available at the wrist joint of the current STS Space Suit, especially during wrist extension could impact on EVA work performance. This condition could influence future EVA tool design; 2) the current EVA knob does not seem to be effective in its design requirement to receive and transfer skeletal muscle force by suited EVA astronauts.

Astronaut testing will continue during 1987. Six zero-g flights are scheduled for February and March. WETF and one-g testing will continue as a comparison to the zero-g testing.

The impact of this study is to provide data in support of STS EVA work requirements, EVA crew training requirements, and Phase C of Space Station. In addition, should the prototype force measuring unit prove successful, the mass and dimension complexities of the Cybex could be eliminated, thereby facilitating future studies.

Specific Underwater Experimental Use of Cybex Dynamometer
==

The attached Hazard Analysis Report describes the safety design

engineering aspects of the Cybex Dynamometer pressurized waterproof

underwater container and electronic safety controls.

The underwater use of the Cybex will be to determine astronaut STS

space suited and unsuited dominant upper limb EVA work related

performance as measured by foot pounds of torque. The following

dominant upper limb movements, with corresponding velocities will

provide the neutrally bouyant (stimulated zero-g) aspect of the data

required for this research.

1) EVA ratchet wrench crank
0°-sec
60°-sec
180°-sec

2) EVA ratchet wrench push and pull
0°-sec
60°-sec
180°-sec

3) EVA knob turn
0°-sec
30°-sec
120°-sec

4) EVA handle turn
0°-sec
30°-sec
120°-sec

5) Wrist flex and extend
0°-sec
60°-sec
180°-sec

6) Elbow flex and extend
$0°$-sec
$60°$-sec
$180°$-sec

7) Shoulder joint flex and extend
$0°$-sec
$60°$-sec
$180°$-sec

8) Forearm rotation (pronation and supination)
$0°$-sec
$60°$-sec
$180°$-sec

9) Shoulder joint horizontal flexion and horizontal extension
$0°$-sec
$60°$-sec
$180°$-sec

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

University of Houston

Expansion of Space Station Diagnostic Capability to
Include Serological Identification of Viral and
Bacterial Infections

Prepared by: Kelly E. Hejtmancik, Ph.D.

Academic Rank: Instructor

University & Department: Galveston College

                                      Division of Mathematics and Science

NASA/JSC

Directorate: Space and Life Sciences

Division: Medical Sciences

Branch: Biomedical Laboratories

JSC Colleague: Duane L. Pierson, Ph.D.

Date: August 8, 1986

Contract #:   NGT44-005-803 (University of Houston)

EXPANSION OF SPACE STATION DIAGNOSTIC CAPABILITY TO INCLUDE SEROLOGICAL IDENTIFICATION OF VIRAL AND BACTERIAL INFECTIONS

Kelly E. Hejtmancik, Ph.D.
Instructor
Division of Mathematics and Science
Galveston College
Galveston, Texas 77550

It is necessary that an adequate microbiology capability be provided as part of the Health Maintenance Facility (HMF) to support expected microbial disease events during long periods of space flight. The applications of morphological and biochemical studies to confirm the presence of certain bacterial and fungal disease agents are currently available and under consideration. This confirmation would be greatly facilitated through employment of serological methods to aid in the identification for not only bacterial and fungal agents, but viruses as well. A number of serological approaches were considered, particularly the use of Enzyme-Linked Immunosorbent Assays (ELISAs), which could be utilized during space flight conditions.

A solid phase, membrane supported ELISA for the detection of Bordetella pertussis was developed to show a potential model system that would meet the HMF requirements and specifications for the future space station. A second model system for the detection of Legionella pneumophila, an expected bacterial disease agent, is currently under investigation. These preliminary studies demonstrate the capability of ELISA systems for identification of expected microbial disease agents as part of the HMF.

INTRODUCTION

The health and well being of individuals aboard a space station and possibly during future long space missions is of priority and must be assured. Certain expected clinical syndromes and diseases have been identified through an infectious disese conference conducted during October, 1985. Previous spaceflight studies indicate a high probability of cross-contamination among crewmembers during long confinements, such as the 90 day missions planned for the space station (12). Continual habitation, crowded conditions, possible immunosuppression, and other factors may create critical situations aboard the space craft. If a microbial disease is suspected, the major effort would be directed toward obtaining some indication of the specific kind of microorganism causing the problem. The exact nature of the etiologic agent would determine the severity of the disease, treatment, prophylaxis, and subsequent health measures for the space station environment.

The diagnosis of a microbial disease rests upon one or a combin- ation of clinical signs and symptoms, morphological and biochemical identification of isolates, and/or serological procedures. Special procedures such as cell culture may also be required. One problem with limiting the scope of diagnosis to clinical signs and symptoms is that a particular microbe can sometimes produce infection having very different clinical characteristics and occuring in widely different areas of the body. For example, antibiotic-resistant Staphylococcus aureus may produce skin and subcutaneous tissue lesions as well as pneumonia, osteomyelitis, bacteremia, and acute membranous entercolitis,

depending upon the means by which the organism gained entrance to the body, host resistance, antibiotic therapy, and other factors.

While the principle of one microorganism causing one clinical disease is often valid, there are many situations where this is not true. Indeed, pneumonias that are hardly separable clinically may be produced by several different kinds of bacteria and viruses. Correct diagnosis and treatment therefore heavily depend upon the abilities of the clinical laboratory.

It is important to note that serological procedures will not immediately take the place of needed morphological and biochemical identification of bacteria or fungi; however, they are frequently used to verify, compare, and further substantiate those results as well as provide a means to directly identify viruses or the immunological response to a viral infection.

Over the past few years, many new immunological methods have been developed which now provide the clinical laboratory with a large array of potentially valuable diagnostic tools. Antibodies and antigens labeled with radioisotopes or fluorescent dyes, or affixed to parti- culate materials, have been used extensively for immunodiagnosis over the past three decades. These methods do have disadvantages. Immuno- fluorescence, for example, usually depends upon a subjective assessment of end result, and the technique is frequently laborious. Radioimmuno- assay requires expensive equipment and carries the risk of radioactive exposure and contamination.

The concepts that antigen and antibody can be attached to a solid

phase support yet retain immunological activity, and that either can
be coupled to an enzyme and the complex retain both immunological and
enzymatic activity, led to the development of Enzyme-Linked Immuno-
sorbent Assays (ELISAs). Antibodies and antigens have been shown to
readily attach to plastic surfaces (such as polyvinyl or polystyrene)
either by passive absorption or chemical conjugation, and still retain
immunological activity. Antibodies and antigens have been linked to
a variety of enzymes including glucose oxidase, peroxidase, and
alkaline phosphatase. The positive factors for ELISAs include low
cost, reagent stability, safety, sensitivity, reproducibility, and
ease of procedure. The procedures are simple enough to be performed
in even poorly equipped laboratories.

It appears likely that the space station diagnostic capability
will require immunological testing applicable to the identification of
microorganisms, particularly viruses. In recent years, there has been
increasing emphasis on accurate, reliable, and quick immunological
procedures for the identification of many microorganisms and/or the
immunological responses of the host toward infection. Most current
procedures have been developed for use in clinical labortories and
not designed for a space station environment. It appears and is
reasonable that a number of some exsisting procedures, particularly
solid phase immunoassays, could be modified in regard to uniformity
and standardization for use aboard the space station. This project
was designed to illustrate the concept of a solid phase, membrane
supported ELISA to demonstrate the capability of ELISA systems for

identification of expected microbial disease agents aboard the space station.

The main purpose of this project was to assess the current ELISA technological trends and procedures in the immunological identification of viral and bacterial diseases, particularly those microorganisms expected to cause illness aboard a space station, and to determine which procedures could be effectively implemented into the space station microbiology diagnostic capability as part of the HMF (Health Maintenance Facility).

MATERIALS

Equipment. A 96 well Bio-Dot filtration apparatus (#170-6550) was obtained from Bio-Rad Laboratories, Richmond, Ca. 94801.

Buffers. A 20 mM Tris buffered saline (TBS), pH 7.5, was prepared by adding 4.84 g Tris (Bio-Rad) to 58.48 g NaCl, brought to a 2.0 l volume with deionized water. The buffer was adjusted to pH 7.5 with HCl.

Blocking Solution. A 3.0% BSA-TBS solution was prepared by adding 3.0 g of bovine serum albumin (Difco) to 100 ml of TBS.

Wash Solution. A wash solution containing 0.05% Tween-20 was prepared by adding 0.5 ml of Tween-20 (Bio-Rad) to 1 l of TBS.

Antibodies. A human serum pool containing antibodies to Bordetella pertussis was obtained from the clinical laboratories at NASA, Johnson Space Center. Antiserum to Legionella pneumophila (serogroup 1) was obtained from Dr. Hazel Wilkinson, the Department of Health and Human Services, Center for Disease Control, Atlanta, Ga. 30333. Horseradish peroxidase conjugated (HPR) goat anti-rabbit (#170-6500) and (HPR) goat anti-human (#172-1050) were obtained from Bio-Rad Laboratories.

Antigens. Bordetella pertussis antigen (#2515) was obtained from Difco Laboratories, Detroit, Mi. 48232. This concentration of this preparation was 2 x McFarlands units (equivalent to approximately 1.8×10^9 organisms/ml). Legionella pneumophila antigen was prepared from a ATCC 3152 (serogroup 1) lyophilized culture vial (13). The ATCC vial was broken and the lyophilized material was dissolved into

4 ml of Trypticase Soy Broth. Four 15 x 100 mm plates containing 25 ml of buffered charcoal yeast extract (BCYE) agar was each inoculated with 1.0 ml of the dissolved material. The plates were enclosed in a plastic container to prevent the agar from drying out and were incubated at 35°C for 48 hr. The cells were suspended from each agar surface in 3.0 ml of sterile 0.01 M phosphate buffered saline, pH 7.2, with a pasteur pipette into a 25 ml sterile conical tube. The conical tube containing the cell suspension was boiled for 1 hr to kill the cells. The killed cell suspension was centrifuged at 1600 x g for 30 min, the supernatant discarded, and the cells resuspended in 2.0 ml of 0.01 M phosphate buffer, pH 7.2 for each 0.1 ml of packed cells. One drop of a 1:1000 methiolate solution was added for each 2.0 ml of preparation. The stock solution was stored at 4°C for 10 days to allow for the release of soluble antigen from the cells. The suspension was centrifuged at 1600 x g and the super- natant used for assay development.

Stock Chromogenic Substrate Stain Solution. Two substrates were utilized for comparison. 0-phenylenediamine (OPD, Abbott Laboratories) was prepared by dissolving 12.8 mg into 5 ml of citrate phosphate buffer, pH 7.2, containing 0.02% hydrogen peroxide. 4-chloro-1-napthol (4ClN, Bio-Rad) was prepared by dissolving 60 mg of 4ClN into 20 ml of ice cold methanol. Immediately prior to use, 0.06 ml of ice cold 30% hydrogen peroxide was added to 100 ml of room temperature TBS. The two solutions were mixed just prior to use.

METHODS

The ELISA systems for both Bordetella pertussis and Legionella pneumonphila utilized the Bio-Dot apparatus with the mounted nitrocellulose paper. The procedure for assembly of the apparatus and preparation of the nitrocellulose paper was provided by Bio-Rad Laboratories (2). Briefly, the nitrocellulose paper was first soaked in TBS to ensure uniform protein binding and low background absorption. The cleaned and dried Bio-Dot apparatus was assembled, and the nitrocellulose paper sheet wetted prior to being placed in the apparatus. The apparatus was appropriately tightened to insure that cross well contamination would not occur.

The flow valve was adjusted to allow the vacuum chamber to be exposed to the atmosphere and the appropriate wells to receive the antigen preparation were inoculated with a 0.05 ml volume. (Proteins bound were minute quantities of either antigen or capture antibody applied as a 0.05 ml volume of a concentration of 0.1-1.0 mg/ml.) Nitrocellulose paper has a protein binding capacity of 0.08-0.1 mg/cm^2 The entire sample was allowed to filter through the membrane by gravity flow (approximtely 30 min). Each well was filled with the same volume of sample solution to insure homogeneous filtration of all sample wells.

After the antigen samples completely drained from the apparatus, 0.2 ml of a 3.0% BSA/TBS blocking solution was applied to each well. Gravity filtration was allowed to occur until the blocking solution completely drained from each well (approximately 30 min).

The flow valve was adjusted to vacuum and 0.4 ml of wash solution (TBS with 0.05% tween) was added to each well. The wash solution was allowed to completely drain from all wells. This process was repeated.

Following the wash step, the flow valve was opened to the atmosphere and 0.1 ml of the first antibody solution was added to each of the wells. The solution was allowed to completely drain from the wells, and another wash step performed.

With the vaccum off and the flow valve to the atmosphere, 0.1 ml of second antibody (HPR antibody against the first antibody) was added to each well. The solution was allowed to completely drain from the wells.

Following the second antibody step, the vaccum was turned on and a wash step performed. Immediately, 0.2 ml of a color development solution, either OPD or 4ClN was applied to each well. A positive ELISA reaction will be shown as color development depending upon the substate utilized.

RESULTS

Non-specific protein binding: A 2.0 cm disc of nitrocellulose paper was appropriately mounted in a modified millipore apparatus. The nitrocellulose was washed twice with TBS. A 2.0 ml aliquot of a BSA solution was allowed to pass through the membrane. A spectrophotometric reading (320 nm) for protein in the solution was taken before and after the solution passed through the membrane. The readings were compared, and it was calculated that the nitrocellulose paper retained approximately 0.110 mg/cm^2 of total protein. This was corrected for the amount of protein released by a wash step.

Enzyme-substrate system: The indicator substrates, OPD and 4ClN, were tested for their interaction to HPR goat anti-human antibody attached to the nitrocellulose paper. Eight rows of 12 cells in the Bio-Dot apparatus were prepared identically, initially washed with TBS, followed by the application of 0.05 ml of serial dilutions ranging from 1:100 to 1:10,000 of the HPR goat anti-human serum. Each cell was blocked against additional protein binding by the coating buffer and washed with TBS. Different volumes of OPD substrate ranging from 0.05 to 0.3 ml, but consistent for each row was applied to the first four rows of the cells. Identical volumes were applied to the last four rows using the 4ClN. Color changes of the substrates were noticable and complete within 5 min. Maximum color change of the OPD substrate occurred with 0.3 ml; however, adequate color change was noticed with 0.2 ml which was subsequently selected for assay development. This colormetric change allowed visualization of membrane attached antibody

to a 1:10,000 dilution. Results with the 4ClN were disappointing.

A purple color change was noted using 0.3 ml of the solution; however,

this occurred with the membrane attached antibody to a 1:300 dilution.

This experiment reflects that 1) the nitrocellulose paper was adequately

binding protein (in the form of antibody), 2) the enzyme-substrate

reaction was appropriate, and 3) the OPD appeared to be superior to 4ClN

for ELISA development.

Bordetella pertussis: A human pool was titrated in the following

manner. Two duplicate rows of cells were prepared in which 0.05 ml of a

1:10 dilution (approximately 1.8×10^8 organisms/ml) of the Bordetella

pertussis antigen was applied to each well with the exception of the

first two. These wells received 0.05 ml of TBS and served as control

wells for the experiment. All wells then received the blocking buffer

and were rinsed with the wash buffer. A human pool was serially diluted

from 1:10 to 1:10,000, and 0.1 ml of each dilution applied to a subse-

quent well. This step was followed by the addition of 0.1 ml of HPR

anti-human serum. Each well was then rinsed with the wash buffer. A

0.2 ml aliquot of OPD was then added to each well. The control wells

showed no color; however, a color change was evident in the antibody

titration wells out to a 1:1000 dilution.

Titration of pertussis antigen. Two duplicate rows of cells were pre-

pared in which 0.05 ml of serial dilutions ranging from a 1:10 dilution

(containing approximately 1.8×10^8 bacterial cells/ml) to a 1:5,000

dilution (containing approximately 3.6×10^5 bacterial cells/ml) with

exception of the first two wells. These wells received 0.05 ml of TBS

and served as control wells for the experiment. All wells the received

the blocking buffer and were rinsed with the wash buffer. A .1 ml

aliquot of a 1:1000 dilution of the human pool was added to each well.

All wells were rinsed with the wash buffer. Each well then received

0.1 ml of HRP goat anti-human antibody diluted 1:3000. The wells were

again rinsed with the wash buffer. All wells then received 0.2 ml of

the OPD solution. Color changes were evident out to a 1:5000 dilution

of the antigen preparation (approximately 1.8×10^5 bacterial cells/ml

or 3.6×10^4 bacterial cells/0.05 ml). These results are summarized in

Table 1.

Bordetella Antigen Dilution	*McFarland Units Applied	Dilution of Human Antiserum			
		1:10	1:100	1:1000	1:5000
1:10	0.02	**4+	2+	1+	0
1:50	0.001	4+	3+	1+	0
1:100	0.002	4+	4+	1+	0
1:500	0.0001	4+	2+	1+	0
1:1000	0.0002	3+	2+	1+	0
1:5000	0.00001	1+	1+	0	0
Control	0	0	0	0	0

*1 McFarland unit is equivalent to approximately 0.9×10^9 bacterial cells/ml.
**Values are expressed as 0 (as comparable to control) to 4+ for comparison of the color intensity of OPD

Table 1. Bordetella pertussis antigen detection.

Legionella pneumophila antigen preparation: The antigen preparation

was subjected to both the Biuret and Lowry protein detection procedures.

The Biuret method showed no detectable protein; however, the results from

the Lowry indicated that the antigen preparation concentration was approx-

imately .025 mg/ml. Subsequent calculations were determined from this estimate.

The legionella antigen preparation was titrated in a manner similar to Bordetella pertussis. Two duplicate rows of cells were prepared in which 0.05 ml of serial 5 and 10 fold dilutions of the preparation were applied. The first well of each row served as controls. All wells received the blocking buffer and were subsequently rinsed with the wash buffer. This step was followed by the addition of 0.05 ml of a 1:1000 dilution of rabbit legionella antiserum. Each well was then washed and inoculated with 0.2 ml of the OPD solution. Color changes were evident out to 250 pg/ml of the antigen prepartion. Results from a typical assay are shown in Table 2.

Legionella Antigen Dilution	Protein Applied (ng/0.05 ml)	Dilution of legionella Antiserum				
		1:10	1:50	1:100	1:1000	1:5000
1:10	25.0	*4+	4+	4+	4+	3+
1:50	5.0	4+	4+	4+	2+	1+
1:100	2.5	3+	3+	2+	1+	1+
1:500	0.5	3+	3+	2+	1+	1+
1:1000	0.25	3+	3+	2+	1+	1+

*Values are expressed as 0 (as comparable to control) to 4+ for comparison of the color intensity of OPD.

Table 2. Legionella pneumophila antigen detection.

DISCUSSION

During the past decade, numerous immunoassays have gained wide acceptance as the methods of choice in the diagnosis of a number of disease states (10). The ideal considerations of a diagnostic test include speed, sensitivity, specificity, accuracy, safety, inexpensive reagents, potential for automation, long reagent shelf life, and broad applicability. Neither immunofluorescence or radioimmunoassay meet all these criteria. Many techniques have been developed recently for the immunological detection of antigens and/or antibodies. Enzyme immunoassays such as the ELISAs are among the most popular both in research (10) and clinical laboratory use for the diagnosis of bacteria, protozoans, and viruses as indicated in Tables 3, 4, and 5, respectively. In general, these tests are user-friendly, reliable, highly

Chlamydia trachomatis	Streptococcus pyogenes
Chlamydelisa (M.A. Bioproducts) *Chlamydiazyme (Abbott)	*TestPack (Abbott) *Ventrescreen (Ventrex) *Quest (Quidel Q)
Mycoplasma pneumoniae	*ICON (Hybritech)
Mycoplasmelisa (M.A. Bioproducts)	
Neisseria gonorrhea	
*Gonozyme (Abbott)	
Salmonella sp.	
*(Kirkegaard and Perry)	

Table 3. Commercially Available Enzyme-Linked Immunosorbent Assays for Bacteria. Asterick (*) denotes antigen detection.

sensitive and specific, and require little time to run. Addition
considerations include that no power source or instruments are required
for the performance of the tests, little equipment is required, and the
reagents used are stable. Positive reactions are contrasted by out-
standing color changes.

Toxoplasma gondii
Toxoelisa (M.A. Bioproducts)
Toxo-G (Abbott)
Toxo-M (Abbott)
Toxostat (M.A. Bioproducts)

Table 4. Commercially Available Enzyme-Linked Immunosorbent Assays for
Protozoans.

The majority of these commercially available ELISA systems are
designed to detect antibody levels in blood plasma or other biological
fluids (i.e. urine) and few have been developed for the detection of
microbial antigens. The Rotazyme (Abbott) and Pathfinder (Kallestad)
kits which detect the presence of rotaviruses in stool specimens
(Table 5); the Chlamydiazyme (Abbott), Gonozyme (Abbott), Salmonella
detection kit (Kirkegaard and Perry), as well as the Test Pack
(Abbott), Ventrescreen (Ventrex), Quest (Quidel Q), and Icon (Hybritech)
for detection of Streptococcus pyogenes in throat swabs are designed
for antigen detection. It appears advantageous to utilize ELISA
systems directed to detect microbial antigens, particularly for the
demonstration of their presence in certain body regions, biological
fluids, or the external environment.

The commercially available ELISA systems were not designed to be

```
| CYTOMEGALOVIRUS                        ROTAVIRUS                          |
|     Cytomegalisa (M.A. Bioproducts)    *Rotazyme (Abbott)                 |
|     CMV-Stat (M.A. Bioproducts)        *Pathfinder (Kallestad)            |
|     Cytomegelisa M (Abbott)                                               |
|     CMV total AB (Abbott)              RUBELLA                            |
|                                            Rubazyme (Abbott)             |
| HTLV III                                   Rubazyme-M (Abbott)           |
|     (Abbott)                               Rubelisa (M.A. Bioproducts)   |
|     (Electro-Nucleonics)                   Rubelisa-M (M.A. Bioproducts) |
|     (Ortho)                                Rubestat (M.A. Bioproducts)   |
|                                                                          |
| HEPATITIS-A ANTIGEN                     RUBEOLA                           |
|     Havab (Abbott)                          Measelisa (M.A. Bioproducts) |
|     Havab-M (Abbott)                                                      |
|                                         VARICELLA                        |
| HEPATITIS-B ANTIGEN                         Varicelisa (M.A. Bioproducts) |
|     Ausab (Abbott)                                                        |
|     Auszyme II (Abbott)                                                   |
|     Corzyme (Abbott)                                                      |
|     Corzyme-M (Abbott)                                                    |
|     (Ortho)                                                              |
|                                                                          |
| HEPATATIS-Be ANTIGEN                                                      |
|     HBe (Abbott)                                                          |
|                                                                          |
| HEPATATIS-DELTA ANTIGEN                                                   |
| .   Anti-Delta (Abbott)                                                   |
|                                                                          |
| HERPES SIMPLEX                                                            |
|     Herpelisa 1 (M.A. Bioproducts)                                        |
|     Herpelisa 2 (M.A. Bioproducts)                                        |
|                                                                          |
| MUMPS                                                                     |
|     Mumpelisa (M.A. Bioproducts)                                          |
```

Table 5. Commercially Available Enzyme-Linked Immunosorbent Assays for
         Viruses. Asterick (*) denotes antigen detection.

utilized in microgravity, and thus, little concern was given to HMF

requirements during their development. However, the Test Pack (Abbott)

released in June, 1986 for purchase, has been tested in the NASA-JSC

laboratory, and its technology appears to be promising for space station

use. This system is solid phase utilizing an antigen capture filter

support, in which fluids are contained through diffusion into an internal absorptive sponge. The system requires approximately 10 minutes to run. The basic flow through system was successfully utilized in zero gravity experiments aboard the KC135.

The most common solid phase supports employed in ELISA systems have been polystyrene microtiter plates (16) and tubes (13) to which either antigen or antibody is passively adsorbed, although other supports such as polystyrene beads (8), sticks (3), and cuvettes (11) have been utilized. Antibodies and antigens have also been passively adsorbed to a number of other supports including polyvinyl (16), polycarbonate (14), aminoalkylsilyl glass (7), and silicone rubber (5). Covalent coupling of antigen or antibody to solid phase supports has been successful using cellulose (15), isothiocynate (4), and polyacrylamide (15). Nitro-cellulose filter paper, used extensively in the development of DNA probe technology due to its ability to bind nucleic acids (1), has been found to nonspecifically bind proteins and has recently been employed as the binding surface on which immunoassays, such as the ELISA, are performed (9).

Results from the experiments conducted in this project and the exsistence of a commercial kit paralleling these findings, provide a current technology to be considered for the HMF. A major advantage to consider with the solid phase filter membrane systems is that the fluids involved in the system can be retained (i.e. little chance of spillage in the space craft environment). The use of solid phase filter supports will be soon expanding and kits will eventually be available

for the identification of those microrganisms, including viruses, expected to cause health problems in the space station environment. Since cell culture is usually required for the identification of viruses, this technology would certainly be an viable alternative.

SELECTED REFERENCES

1. Alwine JC, DT Kemp, GR Stark: Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc Natl Acad Sci 74:5350, 1977.

2. Bio-Dot Microfiltration Apparatus Instruction Manual, Bio-Rad Laboratories, Richmond, Ca., 1984.

3. Felgner, P: A new technique of heterogenous enzyme-linked immuno-sorbent assay stick- ELISA I. Description of the technique. Zbl Bakt Hyg I Abt Org A240:112, 1978.

4. Halbert, SP and M Ankey: Detection of hepatitis B surface antigen (HBsAg) with use of alkaline phosphatase labelled antibody to HBsAg. J Infect Dis, Supplement 136:S318, 1977.

5. Hamaguchi, Y, K Kato, E Ishikawa, K Kobayzski, N Katunuma: Enzyme linked sandwich immunoassay of macromolecular antigens using the rabbit antibody loaded silicone piece as a solid phase. FEBS Lett 69:11, 1976.

6. Jones G. and GA Hebert: "Legionnaires'" the disease, the bacterium and methodology. HEW Publication No (CDC) 79-8375, 1979, p 124.

7. Kato, K, Y Hamaguchi, S Okawa, E Ishikawa, K Kobayashi, N Katunuma: Use of rabbit antibody IgG bound on to plain and aminoaklylsilyl glass surface for the enzyme-linked sandwich assay. J Biochem 82:261, 1977.

8. Miranda, QR, GD Bailey, AS Fraser, HJ Tenoso: Solid phase enzyme immunoassay for herpes simplex virus. J Infect Dis, Supplement, 136:S304, 1977.

9. Monroe, D: The solid phase enzyme-linked immunospot assay: Current and Potential Applications. Biotechniques, May/June, 1985, p 222.

10. O'Beirne, AJ, and HR Cooper: Heterogeneous enzyme immunoassay. J Hist Cyto 27:1148, 1979.

11. Park, H.: Hew technique for solid phase immunoassay: Applications to hepatitis B surface antigen. Clin Chem 25:178, 1979.

12. Pierson, DL: Microbiology support plan for space station. National Aeronautics and Space Administration Publication, JSC No. 32015, 1986.

13. Ruitenberg, EJ, PA Steerenberg, BJM Brose, and J Buys: Reliability of ELISA as control method for the detection of Trichinella spirilis in conventionally raised pigs. J Immunol Methods 10:67, 1976.

14. Smith, KO and WD Gehlebeh: Magnetic transfer devices for use in solid phase radioimmunoassay and enzyme-linked immunosorbent assays. J Infect Dis 136:S329, 1977.

15. van Weemen, BK and AHWM Schuurs: Immunoassay using antibody-enzyme conjugates. Febs Lett 43:215, 1974.

16. Voller, A, DE Bidwell, A Bartlett: Microplate enzyme immunoassays for the immunodiagnosis of virus infections. Manual of Clinical Immunology. Edited by N Rose and H Friedman. Am Soc Microbiol, 1976, pp 506-612.

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

Texas A&M University

Interpreting the Production of ^{26}Al in Antarctic Meteorites

Prepared by:	H. R. Heydegger
Academic Rank:	Professor
University & Department:	Purdue University Calumet Department of Chemistry & Physics

NASA/JSC

Directorate:	Space and Life Sciences
Division:	Solar System Exploration
Branch:	Space Physics
JSC Colleague:	J. E. Keith
Date: 8 August 1986	
Contract #:	NGT-44-005-803 (Texas A&M University)

INTERPRETING THE PRODUCTION OF ^{26}Al IN ANTARCTIC METEORITES

H. R. Heydegger
Professor
Department of Chemistry and Physics
Purdue University Calumet
Hammond IN 46323

Large numbers of meteorites have been concentrated at several locations in Antarctica. Glaciological mechanisms of grossly different time scales ($\sim 10^4$ to $\sim 10^6$ years) have been proposed to account for their transport by the ice, and the frequency distribution of the terrestrial ages of these objects has been suggested as a means of determining the relevant time scale(s). The upper limit to the age of ice in Antarctica which would emerge from such a project is of interest to workers in a variety of other disciplines as well. After a meteorite reaches the Earth's surface, the specific radioactivity of ^{26}Al produced by cosmic rays while it was in space decreases because shielding by the Earth's atmosphere reduces further production to a negligible level. Thus, the known half life of this species can be used to determine the object's terrestrial age if the specific radioactivity at time of fall can be estimated with reasonable accuracy and precision. The several models utilized for these predictions were based on the limited data available nearly two decades ago. In this work we have critically examined the much larger data base now available using multiple parameter regression analyses.

INTRODUCTION

Meteoroids are subject to bombardment by high energy particles while in free space. Such projectiles include both galactic cosmic rays (GCR) and solar particles (SP) which can induce nuclear reactions that result in the transformation of some of the stable nuclei of the target to radioactive product nuclei. After the meteorite reaches the earth's surface, production of the radioactive species is esentially stopped because of the shielding effect of the earth's atmosphere against primary GCR and SP. The specific radioactivity of a given nuclide in a particular portion of a meteorite is dependent upon a number of variables: chemical composition, position in the meteorite with respect to the preatmospheric surface, the primary projectiles' intensity vs. energy spectra time dependence during exposure, etc.

Discovery of the accumulation of large numbers (~5000) of meteorites in ablation zones on the Antarctic ice sheet has lead to interest in using these objects as relict tracers for the mechanism of ice transport. It seems likely that these accumulations result when meteorites which have fallen randomly over the Antarctic surface and were incoporated into and transported with the glacier ice are left behind on the surface as this ice is lost in the ablation zone of the particular sheet. Thus, determination of the time scale for ice movement is possible if the "terrestrial age" i.e. (the time each meteorite has been on earth) can be established. The decay of a radioactive species produced in

space provides a suitable "timer-clock", assuming the amount present at fall can be estimated with reasonable accuracy and precision. Interest has been focused on ^{26}Al because it has a half-life consistent with the the time scales (10^4 - 10^6) proposed for the ice transport from the fall zones to the ablation zones [1-6]. The understanding of time scales in Antarctic glaciology is of interest beyond that discipline. The identification of ice of such great age would provide dated samples for particular ocean sediment and paleoclimatogoly studies as well as for investigation of paleoatmospheric composition [e.g. 7].

Determination of the glaciological mechanism involved for a particular ice sheet would involve:

1. collection of ~$10^{2.5}$ meteorites from the ablation zone;

2. measurement of the current ^{26}Al specific radioactivity (D_o) in each meteorite non-destructively via gamma-gamma coincidence spectrometry to a precision of ~10%;

3. estimation of the ^{26}Al saturation specific radioactivity (D_{oo}) present at fall based on the chemical composition of the object;

4. calculation of the terrestrial age (elapsed time between fall and present) for each meteorite based on present and saturation ^{26}Al values;

5. interpretation of the terrestrial age frequency distribution observed in terms of those expected for postulated transport mechanisms.

This work has been concerned primarily with item 3. In particular, we sought to determine whether published formulae yield D_{oo} estimates sufficiently accurate and precise to permit the time resoltuion in terrestrial ages for required useful conclusions regarding glaciological mechanisms to be drawn.

ESTIMATION OF SATURATION ^{26}Al SPECIFIC RADIOACTIVITY

This problem was first addressed systematically by Fuse and Anders [8] nearly two decades ago. A regression of observed ^{26}Al D_o versus Si, Al, and S content was performed for 34 meteorites assumed to have long exposure ages, with contributions due to Ca and Fe+Ni assumed to be known. Contributions from other elements, including Mg, were assumed to be negligible.

Two years later a different approach was taken by Cressy [9], who used the D_o and elemental composition of eight fractions separated from a single meteorite as the set of observations. The independent variables in Cressy's regression were Mg, Al, and Si, with contributions due to S, Ca, and Fe+Ni assumed to be known.

In 1980, Hampel, et al. [10] used six fractions obtained from three meteorites to derive a third set of coefficients for Mg, Al, and Si, while assuming the values for S, Ca, and Fe+Ni to be known.

Keith and Clark [11] made such an analysis on a set of moon rocks in 1974, but the obvious differences in irradiation conditions (2pi vs. 4pi) and sample surface preservation (atmospheric ablation at fall) cause uncertainty as to the applicability of those results to meteorites.

In order to facilitate comparison of the results of the models cited, each set of coefficients (a_i) has been normalized to yield $a_{Si} = 1$. These results are shown in Table 1, and it is obvious that the three sets based on meteorites are quite disimilar, with the coefficient for such a significant element as Al varying by a factor of 3.

These discrepancies may be due to the small numbers of meteorites considered in two of the studies, to differences in exposure conditions or data selection criteria, and/or to the use of inadequate chemical data. (It might be noted here that cases such as those faced here, where the independent variables show considerable covariance amongst themseves, are particularly prone to yielding biased results from small and/or poor quality data bases.) Therefore, it seems worthwhile to assemble as large a data base as feasible (within a reasonable time) from which to assess the three models proposed. Reported [26]Al specific radioactivities numbering over 500 for 299 non-Antarctic meteorites have been obtained from the literature along with 203 (full or partial) chemical analyses and 165 [21]Ne cosmic ray

exposure ages. Where more than one value for a parameter has been found, the mean value was employed in this study. In a few cases, extreme deviant values were rejected prior to taking the mean. All results reported here were obtained using SAS running under VMS 5.03 on the NASA Johnson Space Center Solar System Exploration Division's VAX 11/780 during the period 19 May to 8 August 1986.

The efficacy of the published models in accounting for the variability in observed ^{26}Al specific radioactivty due to variation in chemical composition was determined in the following manner. For each meteorite of known exposure age (t), the predicted value of the ^{26}Al saturation specific radioactivity was calculated via the prescription for each of the models (D_{pi}), and the observed D_O value was corrected to the saturation value (D_{oo}) as follows:

$$D_{oo} = D_O/(1-T) \quad \text{where } T = \exp(-R) \text{ and } R = t*\ln(2)/t_{26}$$

t_{26} being the known half life of ^{26}Al (0.72 Ma). The extent to which the ratio of D_{pi} to D_{oo} conforms to unity is a measure of the accuracy and precision of the model in predicting the parameter of interest.

Results for the mean value of this ratio over all meteorites for which exposure ages were available in the data base are presented in Table 2, along with their precisions. Although each of the models provides agreement within 15% of the desired D_{oo} value, the large size of the data base provides sufficient precision to confirm that the deviation

from the desired value of unity is significant for each of the models. This indicates the presence of systematic errors. If the principal cause of these discrepancies is variation (or inaccuracy) in chemical compositions, a significant difference in D_{pi}/D_{oo} would be expected among the different classes of meteorites. Mean values for the ratio of interest for meteorites of known exposure age in several major classes are also presented in Table 2, and it is seen that such variation is absent.

In view of the systematic deviations found for predictions from the published models, the recent increase of interest in this problem, and the ready availability of the large data base assembled in this work, it seems worthwhile to perform a new search for a more accurate formula for the prediction of D_{oo}. Such a search was undertaken with quite interesting results. Inverse variance weighted and unweighted regressions of the experimentally derived saturation specific radioactivty values for [26]Al vs. a number of parameters were performed. Presentation of the detailed results of this work is beyond the scope of this report, but the following equation has been found to fit the experimenatal data base with an R-squared of 0.96:

$$D_{oo} = (3.0 \pm 0.5)*Si + (3.6 \pm 1.9)*Al + (0.1 \pm 0.5)*Mg$$

where the chemical symbols stand for the respective elemental abundances in % by weight. This regression was based on 81 cases for which the specific radioactivity of [26]Al, the [21]Ne exposure age, and

C-4

the three elemental abundances were all known.

Since the data base includes finds (i.e. objects identified as
meteorites but which have not been observed to fall), as well as
observed falls, it is worthwhile to see if both subpopulations show
the same systematic deviation. The results shown in Table 3 indicate
that the mean values for the ratio of interest are significantly
different for falls and finds. From this disagreement we infer
that the frequency of finds with D_O significantly less than D_{OO} (i.e.
those unsaturated at fall plus those with terrestrial age greater than
about 0.2 Ma) is greater than the frequency of unsaturated falls
(~8%). Therefore, the inclusion of finds as well as falls would
appear to bias the data base toward lower D_{OO} values . This is an
important conclusion because all of the Antarctic meteorites recovered
to date are finds for most of which there is an absence of measured
exposure ages.

CONCLUSIONS

It has been shown in this work that there is a systematic bias in
estimates of the amount of ^{26}Al expected to be present at fall in a
meteorite of known major element composition when previously published
formulae are employed. The mean specific radioactivity of this nuclide
in finds was also found to be distinguishable from that of falls.
An improved formula for estimating the saturation specific radioactivty

of ^{26}Al expected to be present at fall has been derived from the large data base on non-Antarctic meteorites established for this study. Despite the relatively poor precision yielded by any of the formulae, estimates of this quatity were found to be adequate for use in distinguishing between the principal proposed mechanisms for Antarctic glacier ice transport between the accretion and the ablation zones. Further non-destructive radioactivity measurements in order to establish a large data base for Antarctic meteorites from each of the ice sheets of interest would be a logical next goal.

Table 1. Comparison of Elemental Coefficients in ^{26}Al Estimation Models

Model	Normalized Model Target Coefficient*					
	Mg	Al	Si	S	Ca	Fe+Ni
Fuse & Anders	=0	1.5	=1	.12	=.02	=.007
Cressy	.11	4.6	=1	.54	=.10	=.009
Hampel, et al.	.15	1.8	=1	=.49	=.09	=.011

* = means parameter was set equal to relative value given by original authors and was not a free variable in their regression except the coefficient for Si which was adjusted to unity in this work for ease of comparison.

Table 2. Comparison of Model-predicted Radioactivity by Meteorite Class

Model	Saturation ^{26}Al (Experimental/Predicted) by Class $\pm$ 1 s(m)[*]			
	H	L	C	All
Fuse & Anders	.92 $\pm$ .02 (22)	.90 $\pm$ .04 (15)	.93 $\pm$ .03 (13)	.90 $\pm$ .02 (65)
Cressy	.91 $\pm$ .02 (22)	.89 $\pm$ .04 (15)	.86 $\pm$ .07 (13)	.89 $\pm$ .02 (65)
Hampel, et al.	.91 $\pm$ .02 (22)	.88 $\pm$ .04 (15)	.90 $\pm$ .03 (13)	.88 $\pm$ .02 (66)

[*] inverse variance weighted mean values $\pm$ one sigma of the mean taken over the number of meteorites of known exposure age given in parentheses.

Table 3. Comparison of Model-predicted Saturation Specific Radioactivities

^{26}Al (Experimental/Predicted) Ratio $\pm$ 1 s(m)*

Model	Finds	Falls
Fuse & Anders	.81 $\pm$ .06 (19)	.91 $\pm$ .02 (100)
Cressy	.78 $\pm$ .06 (19)	.85 $\pm$ .02 (100)
Hampel et al.	.78 $\pm$ .05 (21)	.88 $\pm$ .02 (103)

* weighted mean of observed values (uncorrected for nonsaturation) $\pm$ one sigma of mean based on number of meteorites in parenthses.

REFERENCES

1. T. Nagata
   Mem. Nat'l. Inst. Polar Res., Spec. Issue 8, 70 (1979).

2. F. Nishio, N. Azuma, A. Higashi, and J. O. Annexstad
   Ann. Glac.

3. I. M. Willans and W. A. Cassidy
   Science 222, 55 (1983).

4. J. O. Annexstad
   Ph.D. Dissertation, Univ. Mainz, 1983.

5. L. Schultz and J. O. Annexstad
   Smith. Contrib. Earth Sci. 26, 17 (1984).

6. L. Schultz
   Abstracts, Workshop on Antarctic Meteorites, Mainz, 1985.

7. M. Bender, L. D. Labeyrie, D. Raynaud, and C. Lorius
   Nature 318, 349 (1985).

8. K. Fuse and E. Anders
   Geochim. Cosmochim. Acta 33, 653 (1969).

9. P. J. Cressy, Jr.
   Geochim. Cosmochim Acta 35, 1283 (1971).

10. W. Hampel, H. Wanke, H. Hofmeister, B. Spettel, and G. F. Herzog
    Geochim. Cosmochim. Acta 44, 539 (1980).

11. J. E. Keith and R. S. Clark
    Proc. 5th Lun. Conf., Gechim. Cosmochim. Acta, Suppl. 5, 2105
    (1974).

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

University of Houston

Plasma Motor Generator System

Prepared by: Gerald E. Hite

Academic Rank: Associate Professor

University & Department: Texas A&M University at Galveston

                         Department of Marine Science

NASA/JSC

Directorate: Space and Life Sciences

Divsion: Solar System Exploration

Branch   Space Science

JSC Colleague: James E. McCoy

Date: August 1, 1986

PLASMA MOTOR GENERATOR SYSTEM

Gerald E. Hite
Associate Professor
Department of Marine Science
Texas A&M University at Galveston
Galveston, TX 77553

The significant potential advantages of a plasma motor generator system over conventional systems for the generation of electrical power and propulsion for spacecraft in low earth orbits warrants its further investigation. The two main components of such a system are a long (~10km) insulated wire and the plasma generating hollow cathodes needed to maintain electrical contact with the ionosphere.

Results of preliminary theoretical and experimental investigations of this system are presented. The theoretical work involved the equilibrium configurations of the wire and the nature of small oscillations about these equilibrium positions. A particularly interesting result was that two different configurations are allowed when the current is above a critical value.

Experimental investigations were made of the optimal starting and running conditions for the proposed, low current hollow cathodes. Although optimal ranges of temperature, argon pressure and discharge voltage were identified, start-up became progressively more difficult. This supposed depletion or contamination of the emissive surface could be countered by the addition of new emissive material. The sharp transition between the two distinctively different working modes of these hollow cathodes was studied extensively. It is proposed that this transition is due to the establishment in the orifice of a plasma distinct from that inside the hollow cathode.

I. INTRODUCTION

According to Faraday's law of induction, a voltage will be induced in a wire which moves across magnetic field lines. Provided there is a stationary return path for current, electrical power can be extracted (an IxB force will act to reduce the relative speed of the wire with respect to the magnetic field), or with the application of a reversed voltage greater than the induced voltage, electrical power will be expended and propulsion will result. In low Earth orbit a 10-kilometer long wire would have an induced voltage of slightly more than 2 kV. Neglecting losses and assuming good electrical contact with the ionosphere, 20 kw of power would be equivalent to a propulsion thrust of about 2.5 N. [1]

If there is no current in the wire and the spacecraft is in a circular orbit, then the combined effects of gravity and centripetal acceleration will cause the wire to hang either straight down or vertically upward. However, when a current flows, the IxB force will pull the wire from the vertical direction and there will be one, two, or no stable configuration, depending on the magnitude of the current. Clearly, knowledge of these configurations, and of the nature of small oscillations about these configurations, is essential to ensure the integrity of such power/propulsion systems.

The essential elements in the proposed system is the establishment of good electrical contact with the ionosphere. The most convenient way to establish this contact is with a plasma generator capable of producing plasma in such quantities

that it will exceed the ionosphere plasma out to a distance of ten meters or so. Hollow cathodes are known to be copious producers of plasma and have the potential to act as contractors.[2-5] Electrical power/propulsion systems using plasma contactors are generally called Plasma Motor/Generators (PMG).

In the past, hollow cathodes have served as plasma generators in Kaufman thrusters and similar ion propulsion systems. [6] In contrast to previous applications which needed massive ions (e.g., Hg) for propulsion, in this application the emphasis is on the quantity of plasma and thus harmless inert gases suffice as plasma stock. The present need is for low discharge current (~1A), high gas-flow (30 st. cc/min.) hollow cathodes. It is anticipated that little or none of the extensive magnetic field hardware associated with Kaufman thrusters will be necessary. Considerable research is now being done to test the performance of rugged, striped down hollow cathodes for a PMG system. Preliminary experimental testing has demonstrated that two plasma generators running at opposite ends of a vacuum chamber can generate plasmas sufficient to carry several amps of current between them. [7]

The internal dynamics of hollow cathodes are at best only partially understood. [2-5,8,9] At least two distinct running modes are described in the literature. [4,5,9] At low gas flow, hollow cathodes running in the "Plume" mode are characterized by an illuminous plume emanating from the cathode. With increasing gas flow, a value of the internal pressure is reached at which the plume is replaced suddenly by a glowing spot in the orifice

of the cathode, and the discharge voltage required to maintain constant discharge current jumps to a much lower value. This later, low voltage, high gas flow in general, more stable mode (Spot mode) is the preferred mode for a PMG contactor.

Clearly, optional starting conditions for these newer hollow cathodes need to be established to ensure the reliability of a PGM system. There is a long-standing problem of deterioration of hollow cathodes exposed to moisture and other contaminators normally avoided in laboratory environments. Care must be taken to ensure the purity and concentration of the emissive material in the hollow cathodes used in PMG systems.

This report will concentrate on theoretical results concerning configurations and small oscillations of PMG tethers and on experimental studies of optimal starting and running conditions of the cathodes especially designed and constructed for a PMG system by James E. McCoy at NASA Johnson Space Center.

II. CONFIGURATIONS OF AN ELECTRODYNAMIC TETHER

The equilibrium configuration of an electrodynamic tether can be found by considering the forces on an arbitrary, infinitesimal, segment of the tether. It will be convenient to consider a coordinate system moving with the spacecraft which will be assumed to be in a circular, low-earth orbit and to use the following symbols:

    B = component of magnetic field cut by tether
    g = acceleration due to gravity at orbit
    I = current in tether

L = length of tether

R = radius of orbit

T(x) = tension in tether

Ts = tension due to possible subsatellite on end of tether

x = radial distance measured from spacecraft

X = projection of tether on x-axis

θ(x) = angle of tether measured with respect to x-axis

ω = angular speed of the spacecraft

μ = mass per unit length of the tether

The sum of forces acting on such a segment, i.e., drag force, tensions on its ends, gravitational force, and the electrmagnetic force, IxB, must equal the mass times the acceleration. (The drag force can be neglected for the thin wires proposed for the tether.) The acceleration will be separated into a centripetal term and an acceleration with respect to the spacecraft. The centripetal term can be subtracted from the gravitational force to give a "force" which vanishes at the spacecraft.

Assuming a simple fall off of the gravitational force and using the coordinates system shown in Figure 1, the effect of gravity and centripetal acceleration for a mass, m, of tether is:

$$F_g = mg\,(R/r)^2 - m\omega^2\,r \qquad , \; r = R - x$$

(1) $$\qquad = 3mg\left(\frac{x}{R}\right) + O\left(\left(\frac{x}{R}\right)^2\right)$$

The equations of motion for a piece of tether of mass, m, and

length, l, are:

$$(2a) \quad F_g + T_R \cos\theta_R - T_L \cos\theta_L - IB\ell \sin\theta = m\ddot{x}$$

$$(2b) \quad IB\ell \cos\theta + T_R \sin\theta_R - T_L \sin\theta_L = m\ddot{y}$$

where the subscripts R and L indicate the values at the right and left ends of the segment shown in Figure 1.

FIG. 1
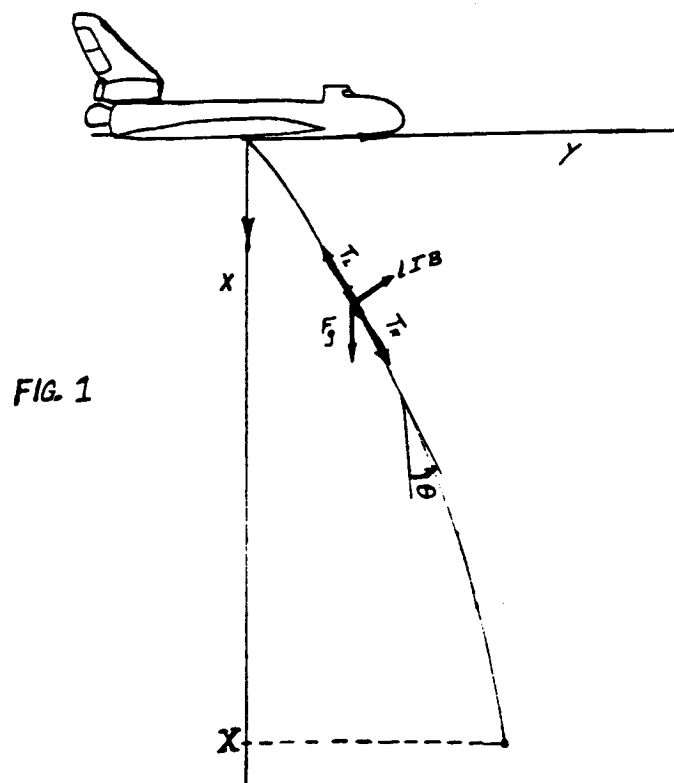
Expanding equation (2) in $\Delta x (= l \cos \theta)$ and retaining first order terms as Δx goes to zero yields:

(3a) $$3 \mu g \left(\frac{X}{R}\right) + \partial_x T = \mu \left(\ddot{y} \sin \theta + \dot{x} \cos \theta\right) / \cos \theta = \mu \ddot{x}_{\parallel} / \cos \theta$$

(3b) $$I B + \partial_x (T \sin \theta) = \mu \ddot{y} / \cos \theta$$

where $\ddot{x}_{\parallel}$ is the acceleration in the direction of the tether.

The static configurations are obtained by setting the accelerations in equation (3) equal to zero. The first equation yields the tension:

(4a) $$T(x) = \frac{3}{2} \left(\frac{\mu g}{R}\right)\left(X^2 - x^2\right) + T_s$$

while the second equation gives:

(4b) $$T \sin \theta = I B \cdot (X - x)$$

Clearly, the tension or $\sin \theta$ vanish at the far end of the tether. In the case when the subsatellite has negligible mass equation (4) yields:

(5) $$\sin \theta = a / (X + x)$$

where $\quad a/L = (2BR/3Mg) \cdot I = I/(\sqrt{3}\, I_c)$

The projection of the tether on the x-axis, X, can be related to the tether length, L, by:

(6)
$$L = \int_0^X dx/\cos\theta$$

In the absence of a subsatellite, this integral gives:

(7)
$$L/a = \sqrt{\left(\frac{2X}{a}\right)^2 - 1} - \sqrt{\left(\frac{X}{a}\right)^2 - 1}$$

FIG. 2
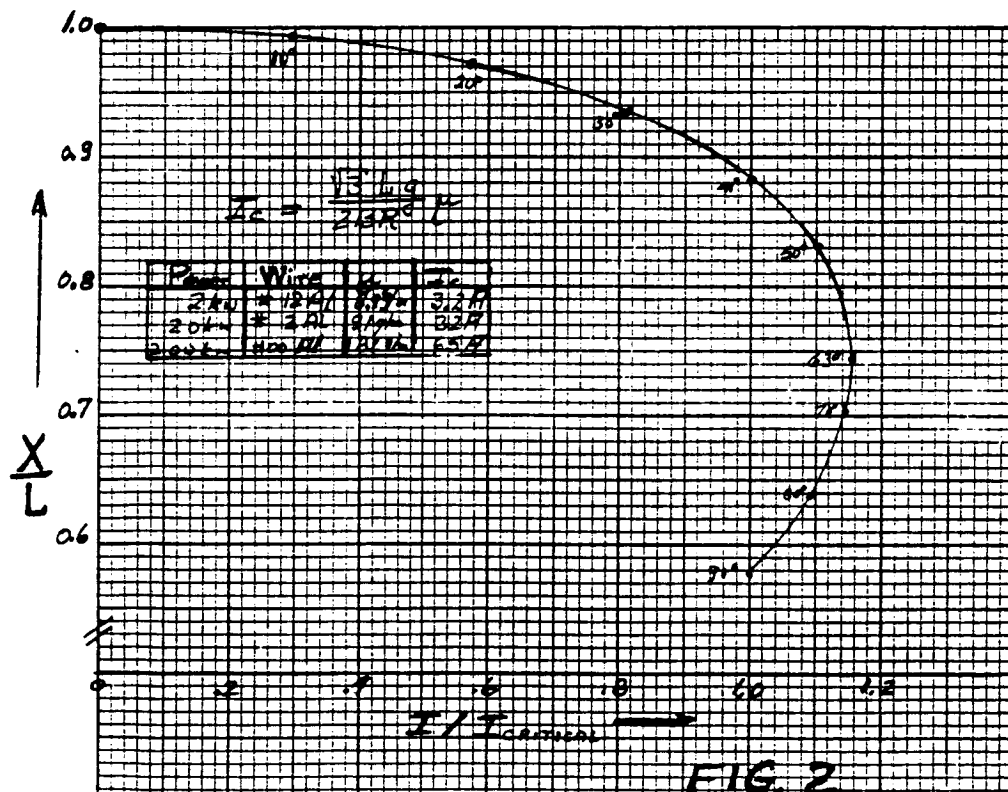
A plot of I/I_c versus X/L is shown in Figure 2. For values of current less than I_c, there is one value of X and thus one stable configuration. For the current between I_c and $I_c \cdot \frac{2}{\sqrt{3}}$, there are two possible values of X and thus two distinct configurations. For still larger currents, the tether does not have a stable configuration and rotates around the spacecraft. Also shown in Figure 2 are the values of θ at the point of attachment on the spacecraft for each configuration.

Although no analytic solution for X was found for nonzero values of T_s, it is reasonable to assume that a small subsatellite would only effect the far end of the tether and that the essential features of the above discussion would still be true.

The nature of small oscillations about the stable configurations can be obtained from equation (3). For small oscillations the tether will be essentially only displaced perpendicular to its stable position, i.e., the right side of equation (3a) is zero and the tension will still be given by equation (4a). The normal procedure is now to replace $\sin\theta$ by $\tan\theta = \partial_x y$ which is the small angle approximation. Since for the tethers considered the current will be much smaller than I_c, the small angle approximation will be acceptable. Using this replacement and letting $\tilde{y}$ be the displacement away from the stable configuration, equation (3b) reduces to:

$$(8) \qquad \ddot{\tilde{y}} = \frac{1}{t} \partial_x (T \partial_x \tilde{y}) = \frac{3}{2}\left(\frac{gX^2}{R}\right) \partial_x \left(1 - \frac{x^2}{X^2}\right) \partial_x \tilde{y}$$

Assuming harmonic solutions, the solution for $\tilde{y}$ is given by Legendre polynomials of odd order, i.e.,

$$(9a) \qquad \tilde{y} = \tilde{y}_o \cos \omega_\ell t \cdot P_\ell \left(\frac{x}{X}\right) \quad , \quad \ell = 1, 3, 5, \cdots$$

where

$$(9b) \qquad \omega_\ell^2 = \frac{3}{2} \left(\frac{g}{R}\right) \ell(\ell+1)$$

The frequency of small oscillation is just the square root of an *integer* times the orbital angular frequency. The excitation energy is given by:

$$(9c) \qquad \mathcal{E}_\ell = \int_o^X \frac{1}{2} \mu (\omega \tilde{y}_o)^2 dx = 3 \left(\frac{M_g}{R}\right)\left(\frac{X}{L}\right) \tilde{y}_o^2 \frac{\ell(\ell+1)}{2\ell+1}$$

The amplitude $\tilde{y}_o$ is the maximum displacement at the far end of the tether from equilibrium.

The above results is exact for small values of the current, e.g., a fraction of I_c. For values up to I_c equations (9) should still give approximate results. The existence of two configurations for I larger than I_c would imply that if sufficient oscillatory energy were available, the tether would swing over a wide region containing both configurations.

All such oscillations could be reduced by applying an AC voltage to the tether with the resonance frequency but lagging by 90°. This would be just the opposite situation to that of a forced oscillator where the driving force leads by 90° at resonance. A more complicated means of dampening oscillations involves changing the length of the tether.

III. The Plasma Generator

The required electrical contact with the ionosphere for a PMG system could be maintained through the plasma generated by a low power, high gas flow, rugged hollow cathode running on a noble gas such as argon. Fig. 3 shows a cross-sectional veiw of the proposed hollow cathode.

FIG 3
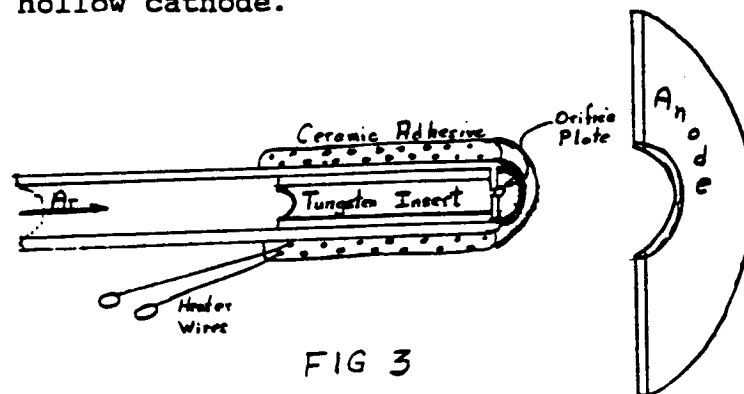
The external heater wires heat the cathode to over 1000°C. The porous tungsten insert is originally impregnated with $BaCO_3$ and $SrCO_3$. The gas carrying tube is made from tantalum, the orifice plate of tungsten, and the anode of molybdenum.

It is believed that the initial heating converts the $BaCO_3$ and $SrCO_3$ into BaO and SrO. Such alkaline earth compounds have very low work functions and are able at high temperatures to thermally (with possible assistance from sheath induced electric fields) emit electrons.[5] The electrons accelerated by electric fields obtain sufficient energy to ionize the neutral gas atoms either by single or multiple step processes.

The proposed heater is designed to be simple and rugged. The heating wire is 90° tungsten-10% shodium thermal couple wire. Consequently, the variation of its resistivity with temperature

is known. By measuring the resistance of the heater at each voltage setting, i.e., V_{heater}/I_{heater}, and normalizing by the value at room temperature, it is possible to determine the temperature of the heating wire for each voltage setting. As shown in Fig. 4, the temperature corresponding to the normal heating voltage (28v) is abut 1550°C. Clearly a large percentage of the heat is being radiated outward and lost. A tantalum foil jacket could be put around the heater to conserve heat, but it was considered to be an unnecessary complication that could reduce the sturdiness of the plasma generator.

The small orifice (15 mil. dia) in the orifice plate enables the hollow cathode to have a relatively high internal pressure (150mm) appropriate to the optimal running conditions at the high gas flow normally used (30 st. cc/min).

Optimal starting conditions were established for the four identical hollow cathodes tested. It was found that a heating voltage of 28v (i.e., ca. 80 watts), and an internal pressure of approximately 130mm would allow a fresh hollow cathode to start spontaneously in the desired mode of operation (spot mode). However, after several starts and/or an accumulated running time (on the order of a few hours), the hollow cathodes became progressively harder to start and eventually refused to start spontaneously. Difficult starts were accompanied with large flashing plumes, spitting of "sparks" from the orifice, etc. Once running, no apparent deviation from normal operation was observed. Todate, it has not been possible to find the cause of this deterioration. Similar degradation of hollow cathodes using

FIG 4
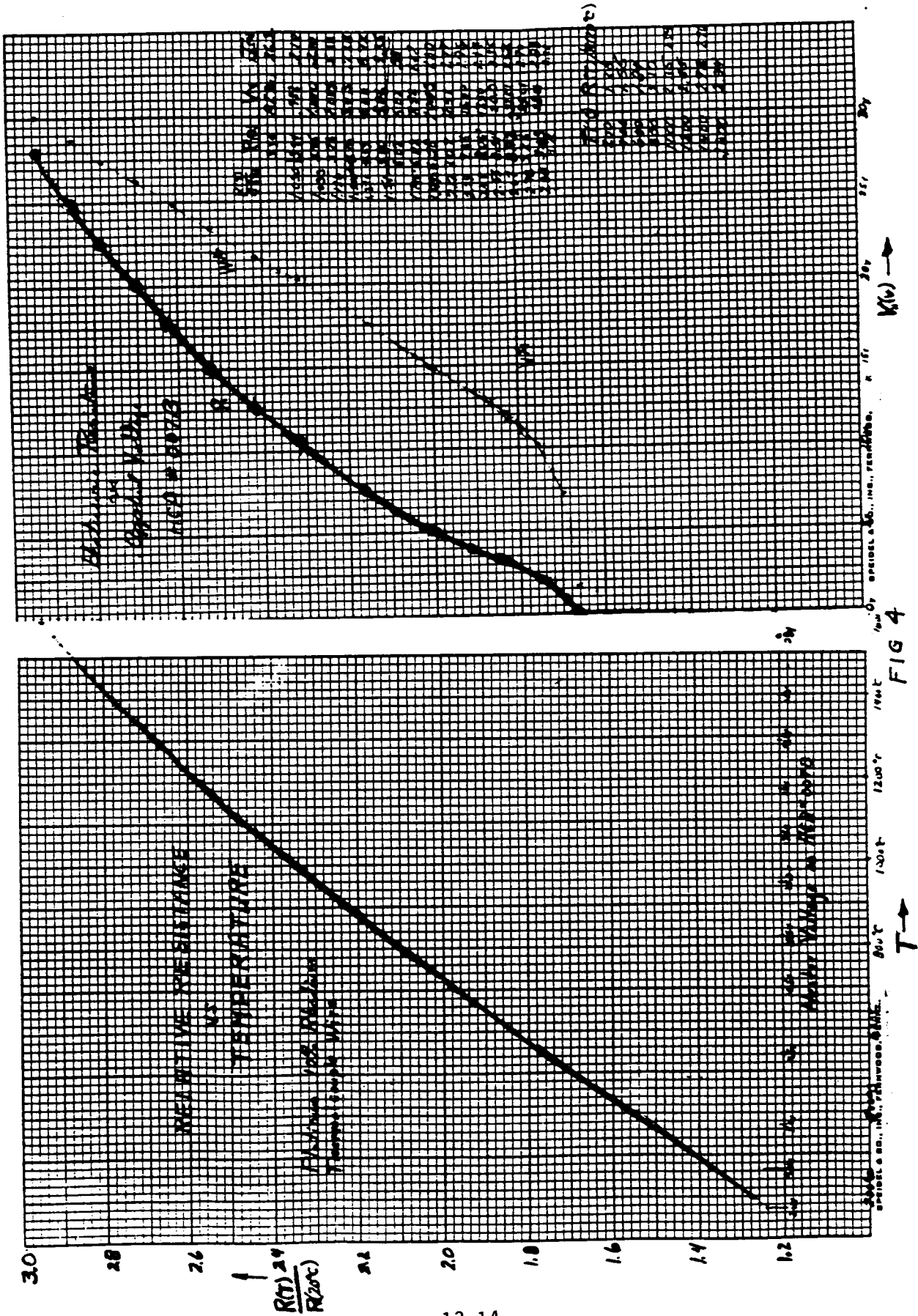
alkaline earth oxides has been observed and studied. [10] This supposed depletion or contamination of the emissive surface could be countered by the addition of new emission material. In some cases, starting could be initiated by increasing the internal gas pressure to over 200mm. Since the vacuum system could not handle such flow rates for long, reliability of this method was not pursued. Starting was possible in nearly all cases, if a 3 kv spark located near the anode was excited two or three times. Similar starts were induced by turning on an ion-vacuum gauge or a second hollow cathode elsehwere in the chamber. Apparently, the ions of such plasmas would travel to the cathode and affect the emission conditions or possible space charge configurations sufficiently to allow the hollow cathode to start.

A model of how externally produced ions could initiate emission has been proposed. [7] Such ions would impinge on the outer surface of the orifice plate and initiate emission there. Assuming that emissive material migrates from the insert inside the cathode, through the orifice and onto the orifice plate, such emissive regions would tend to migrate to regions of increasing concentrations of emissive material, which would be in the direction of the orifice. Consequently, the emissive region and its accompanying and growing sheath would migrate into the orifice and, if conditions are favorable, into the interior of the cathode. Support for this explanation comes from the fact that a high voltage, low discharge (mA), is often observed before ignition even with no gas flow. During such discharges, a small (2-3 mm long) plume is often seen coming off the orifice plate.

In one case, it was at the position where emissive material had
leaked out of the orifice after such material had been injected
into the interior of the cathode.

Extensive investigations were made of the optimal running
conditions for the PMG hollow cathodes. In general, most hollow
cathodes have two distince modes of operations. [2-5, 7-10] At low
gas flow rates and internal pressure, hollow cathodes usually are
running in the so-called plume mode, in which a blue plume is
seen emanating from the orifice and extending to, or in some
cases, past the anode. The shape and position of the anode
influences the plume. As the external gas pressure is increased,
a characteristic, transition value is reached at which the plume
vanishes and leaves a bright glowing spot in the orifice (hence
the name spot mode). Whereas before the transition the discharge
voltage required for constant discharge current is decreasing,
during the transition it falls abruptly to half of its value
before the transition. The data shown in Fig. 5 shows the
transition as the gas pressure is decreased. Although the spot
mode is more stable mode than the plume mode, the instability of
the plume mode in Fig. 5 is accented by gas pressure being
changed too fast to allow the system to obtain complete thermal
equilibrium.

The spot mode is a high gas flow, low power consumption
mode. However, measurements have shown that the ratio of ion
production to power consumption appears to be the same for both
modes for other hollow cathodes. [7] Undoubtedly more
experimental investigations must be made on the plasmas produced

GAS PRESSURE   P(mm) ➝
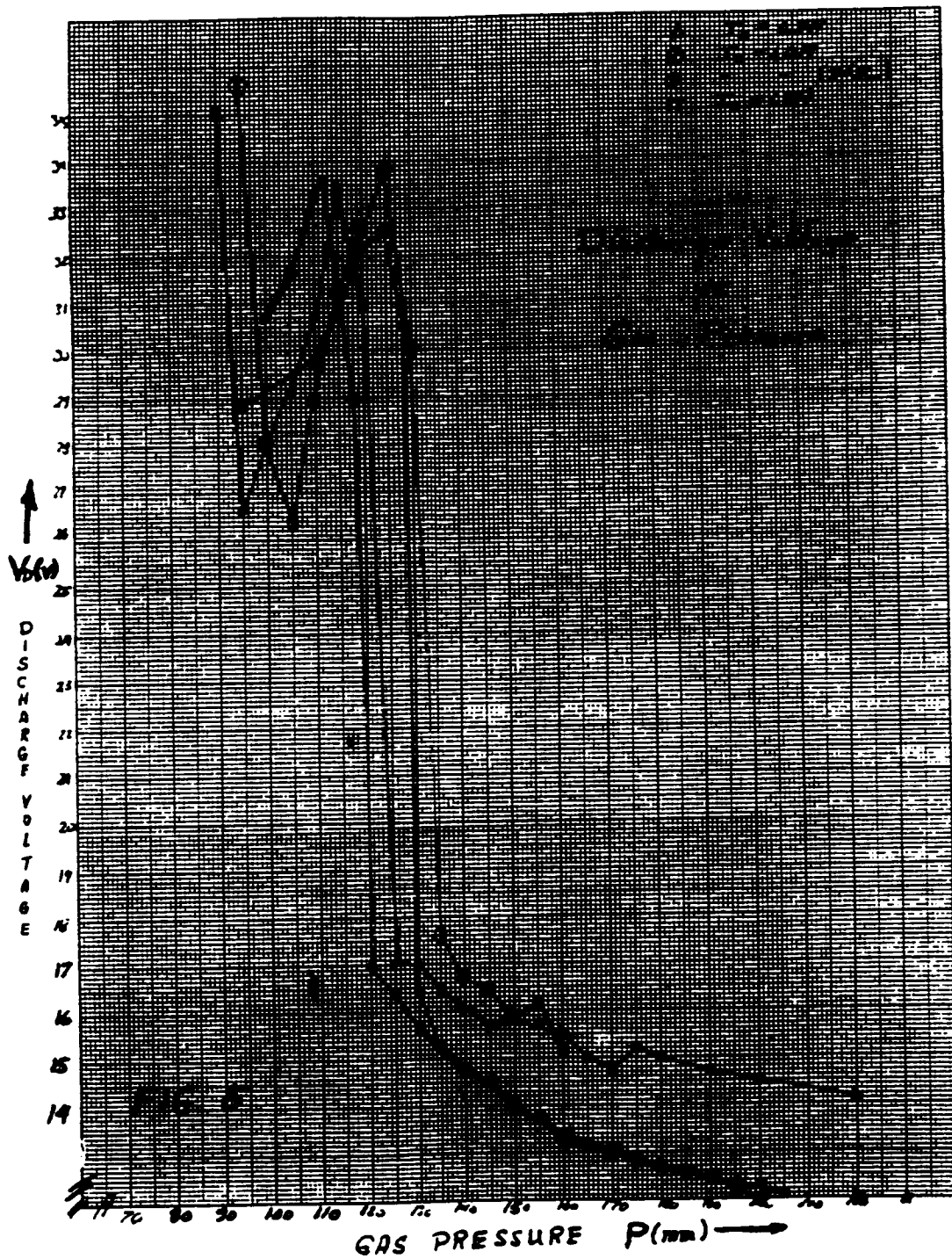
in the two modes before a final decision can be made on the optimal running mode for the PMG hollow cathode.

Extensive modeling of the internal dynamics of hollow cathodes running at low gas flow rates has been done. [2,5,8] Measurements of potentials have indicated that a potential of 8-12 volts is maintained inside the cathode just behind the orifice plate by an ion sheath surrounding a plasma in that region. [5] The upstream extent of the plasma was on the order of a few millimiters and was observed to decrease with increasing pressure. Oscillations of the discharge currents have been observed at presure just above the transition pressure and attributed to oscillations of a sheath in the orifice. [11] It is known that such sheaths are formed when electrons are confined to flow through a construction or orifice. [12] Since the hollow cathodes used in this investigation have run at discharge voltages as low as 5 or 6 volts in the spot mode, it is reasonable to speculate that the plasma in the tube during the plume mode is compressed into a shorter and shorter cylinder and is abruptly forced out of the gas tube and into the orifice. As discussed before, it is reasonable to assume that emissive material has migrated into the orifice and consequently emissions could easily take place from the walls of the orifice. Due to the higher densities in the orifice, the plasma there would be distinctively different from that in the gas tube during plume mode. It is possible that the formation of the orifice plasma and the subsequent emission in the orifice characterizes the onset of the spot mode without complete cessation of emissions

inside the gas tube. The appearance of the bright spot in the orifice characterizing the spot mode could be due to dynamical effects taking place in the downstream boundary of the plasma in the orifice.

In conclusion, it is conjectured that the transition from plume to spot mode is characterized by the formation of an orifice plasma and subsequent field-enhanced thermionic emission from the orifice.

IV. <u>Conclusions</u>

A brief description of the proposed Plasma Motor/Generator system was given. Special attention was paid to the configuration and oscillations of the wire tether and to the starting and running conditions of the hollow cathodes designed to produce the plasma necessary for electrical contact with the ionosphere.

In studying the tether configurations it was found that there could be one, two, or none depending on the strength of the current. Small oscillations about the configurations anticipated in the proposed system were studied. Their form, frequencies and energies were found. Of the two methods of dampening such oscillations, the use of an AC source tuned to the resonance frequency was preferred.

It was concluded from experimental tests that starting of the PMG hollow cathode could be ensured by either frequent introductions of new emissive material or the use of a device capable of geneating a seed plasma such as a sparker. A model

was proposed to explain the dynamical difference between the two distinct running modes observed in the PMG hollow cathode.

References

1. McCoy, J.E., "Electrodynamic Tethers, I. Power Generation in LEO, II. Thrust for Propulsion & Power Storage". Intern Astronautical Federation, 35 Congress, Oct 1984, Paper 440.

2. Krishnan M., John R.G., von jaskowsky W.F., Clark K.E. "Physical Processes in Hollow Cathodes". AIAA Journal Vol 15 No. 9, Sept 1977, pp 1217-1225.

3. Siegfried D.E., and Wilbur P.J., "Studies on an Experimental Quartz Tube Hollow Cathode, Electric Propulsion and its Applications to Space Missions, Vol 79 of Progress in Astronautics and Aeronautics, 1981, pp 262-277.

4. Siegfried D. E., and Wilbur P.J., "an Investigation of Mercury Hollow Cathode Phenomena, 13th Intern. Electric Propulsion Conf. April 1978.

5. Siegfried D.E., A Phenomenological Model for Orified Hollow Cathodes, NASA CR 168026 Dec. 1982.

6. Wilbur P.J. and Kaufman H.R., "Double Ion Productions in Argon and Xenon Ion Thrusters" Journal of Spacecraft and Rockets, Vol 16, No. 4, July-Aug 1979, pp 264-267 and references therein.

7. Wilbur P.J., Colorado State University, Private Communication.

8. Fearn D.G., and Philip C.M., "An Investigation of Physical Processes in a Hollow Cathode Discharge" AIAR 9th Electric Propulsion Conf. April, 1972.

9. Philip C.M., "A Study of Hollow Cathode Discharge

Characteristics", AIAA 8th Electric Propulsion Conf.,
Aug-Sept, 1970.

10. Zuccaro D., "Mercury Vapor Hollow Cathode Component Studies"
AIAA 10th Electric Propulsion Conf. Oct-Nov 1973, Paper No.
73-1141.

11. Siegfried D., Hollow Cathode Plasma Oscillations
ion & Adv. Electric Thruster Research, Dec 1980, pp. 32-29,
NASA CR-165253.

12. Crawford F.W. and Freeston I.L., "The Double Sheath at
Discharge Construction", Proc. Intern. Conf. Phenomena
ionized Gases, 6th, Paris, Vol 1, July 1963, pp. 461-464.

Crawford F.W. and Lawson J.L., "Some Measurements of
Fluctuations in a Plasma", Journal of Nuclear Energy Part C,
Plasma Physics Vol 3, 1961, pp. 179-185, Also Phys. Rev.
Letters, Vol 6. No. 12, June 15, 1961, pp. 663-667.

1986

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center

University of Houston

A Comparison of Two Conformal Mapping Techniques
Applied to an Aerobrake Body

Prepared by: Dr. Mark J. Hommel, P.E.

Academic Rank: Associate Professor

University   Prairie View A&M University
& Department: Department of Mechanical
                           Engineering

NASA/JSC
Directorate: Engineering

Division: Advanced Programs Office

Branch: Aerosciences

JSC Colleague: Dr. C.P. Li

Date: August 8, 1986

A COMPARISON OF TWO CONFORMAL MAPPING TECHNIQUES APPLIED TO AN AEROBRAKE BODY

Conformal mapping is a classical technique which has been utilized for solving problems in aerodynamics and hydrodynamics for many years. Conformal mapping has been successfully applied in the construction of grids around airfoils, engine inlets and other aircraft configurations. These shapes are transformed onto a near-circle image for which the equations of fluid motion are discretized on the mapped plane and solved numerically by utilizing the appropriate techniques. In comparison to other grid-generation techniques such as algebraic or differential type, conformal mapping offers an analytical and accurate form even if the grid deformation is large. One of the most appealing features is that the grid can be constrained to remain orthogonal to the body after the transformation. Hence, the grid is suitable for analyzing the supersonic flow past a blunt object. The associated shock as a coordinate surface adjusts its position in the course of computation until convegence is reached.

In the present study, conformal mapping techniques have been applied to an Aerobrake Body having an axis of symmetry. Two different approaches have been utilized:
    (1) Karman-Trefftz Transformation
    (2) Point-Wise Schwarz-Christoffel Transformation
In both cases, the Aerobrake Body was mapped onto a near-circle, and a grid was generated in the mapped plane. The mapped body and grid were then mapped back into physical space and the properties of the associated grids were examined. Advantages and disadvantages of both approaches were discerned.

1. Introduction

A problem of interest to NASA involves the hypersonic flow past an aerobraking orbital transfer vehicle. As summarized by Li (1), several schemes have been utilized to simplify the numerical treatment of the problem. A primary simplification involves the mapping of the characteristic mushroom shape of the aerobrake vehicle onto a near-circle, generating a grid, solving the Navier-Stokes equations in the mapped plane, and subsequently mapping the solution back into physical space.

In examining the features of the grids generated by this procedure, competitive alternative methods have been re-discovered from elementary complex number theory. The two complementary methods of interest in the present study are as follows:

(1) Karman-Trefftz Transformation

(2) Point-wise Schwarz-Christoffel Transformation.

In the following brief report, both transformations are examined with respect to their suitability for transforming the aerobrake vehicle to a shape suitable for an existing Navier-Stokes computer code, and conclusions and recommendations regarding the two transformations are made. Finally, in order to gain deeper insight into the transformations, a simple square is also transformed by both methods, and the resulting grids examined as well.

2. Karman-Trefftz Transformation

The Karman-Trefftz transformation, which maps the z-plane into the w-plane, is given by the following relationship:

$$\frac{w-1}{w+1} = \left(\frac{z-h}{z+h^*}\right)^{\delta} \tag{1}$$

where z represents the complex physical plane x+iy; w represents the complex mapped plane u+iv; δ is a real number to be defined below; and h is a "hinge point," which is a point in the vicinity of some point of interest on the body being transformed. The transformation has the property of smoothing out corners on the physical body, facilitating the generation of a grid around the body. Each sharp corner on the body is smoothed out in turn by repeated applications of the current transformation to every point on the body, each transfromation having a specific value of h and δ . The real number δ is evaluated as follows:

$$\delta = \frac{\pi}{2\pi - \alpha_{int}} \tag{2}$$

where α_{int} is the interior angle formed at a given corner of the physical body. Following Moretti (2), repeated applications of the transformation are applied in order of increasing δ . In the following figures, transformations are applied to a mushroom-shaped Aerobrake Body. Note in particular the values for h and δ for each transformation. Also note that the actual grid is generated in the mapped plane, where the body has come to resemble a near-circle, by constructing equiangularly spaced radial lines and their orthogonal complements. The grid and body are then mapped back into physical space by reversing the transformations. The computer program for this process is listed in Appendix 1.

Figure 1 shows the initial mushroom configuration, with the sharp corners numbered in the order to be transformed. Note that the first angle to be transformed is the perpendicular angle at point 1, so the exponent in the transformation is 2/3. The result of this first transformation is given in Figure 2, where the the sharp angles at points 2 and 3 are seen to persist, but the corner at point 1 has now been smoothed out. Again the perpendicular corner at 2 is chosen, so the exponent is 2/3, and the result of this transformation is shown in Figure 3. Finally, the corner at point 3 is transformed by using an exponent of 2, and the near-circle of Figure 4 is obtained.

$$\alpha_1 = \frac{\pi}{2} \qquad \delta_1 = {}^2\!/_3$$
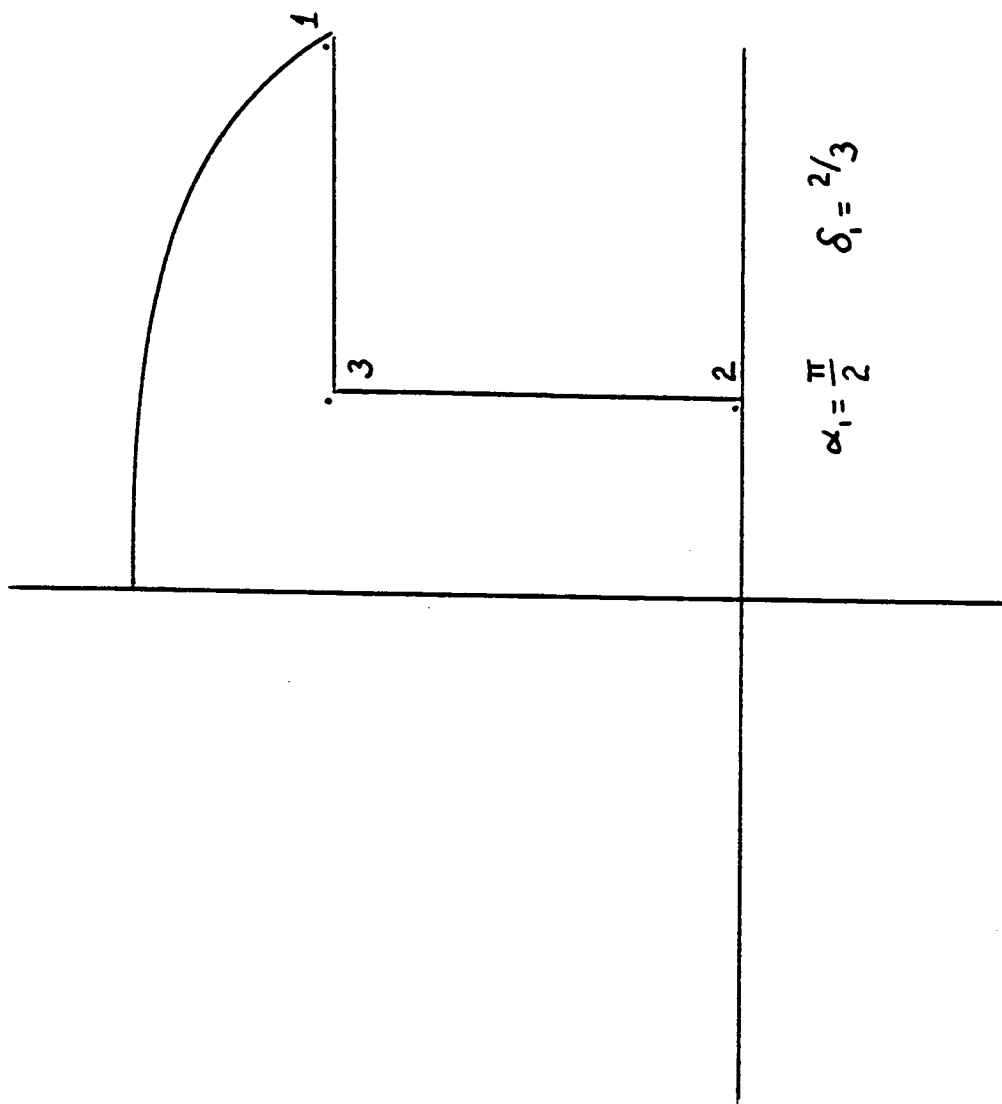
Figure 1--Mushroom Configuration for Karman-Trefftz Transformation

$$\alpha_2 = \frac{\pi}{2} \quad \delta_2 = 2/3$$
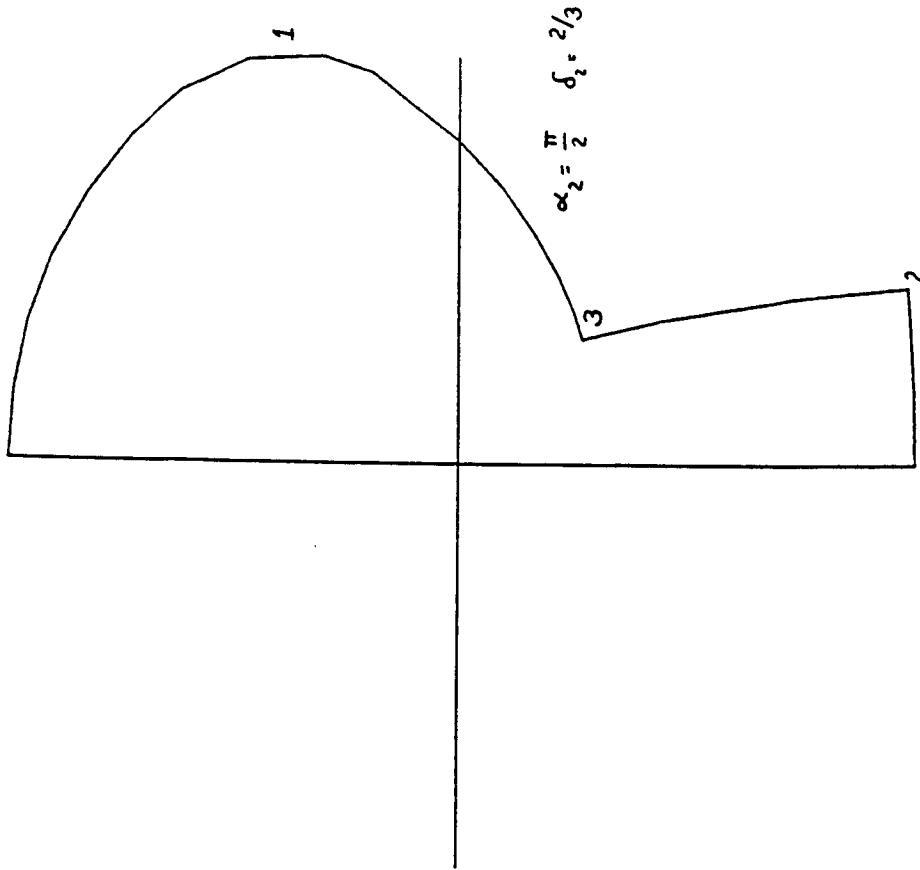
Figure 2--Result of First Transformation
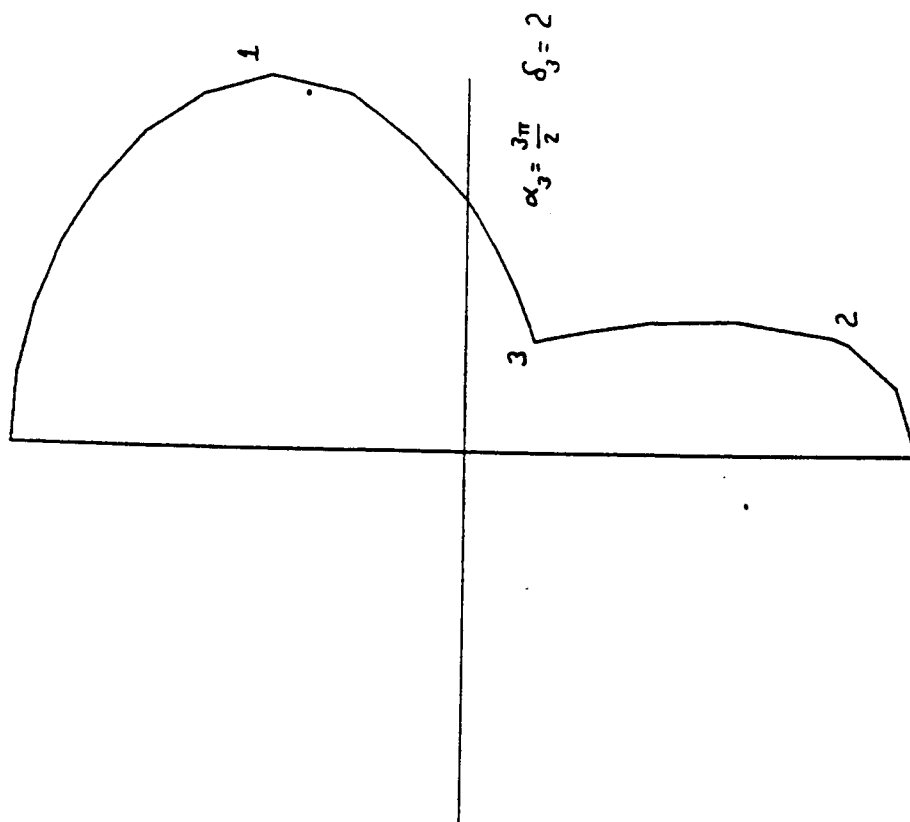
Figure 3--Result of Second Transformation
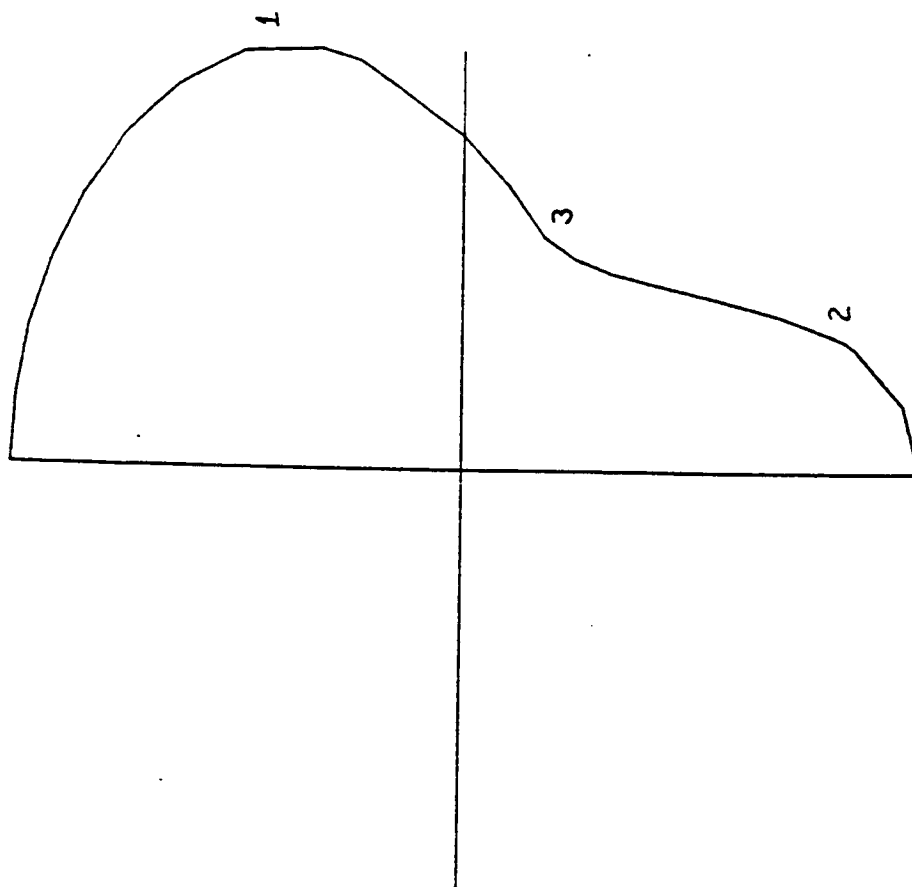
Figure 4--Near-Circle

Next, equi-angular radial lines are constructed from the origin, as well as their orthogonal complements, as shown in Figure 5. It is this set of points which are mapped back into the physical plane to become the grid in physical space. At this point, moreover, it is envisioned that the flow past the Aerobrake body could be obtained numerically, utilizing an existing Navier-Stokes computer code. Figures 6-8 show the resulting mesh in physical at successively improved levels of refinement.
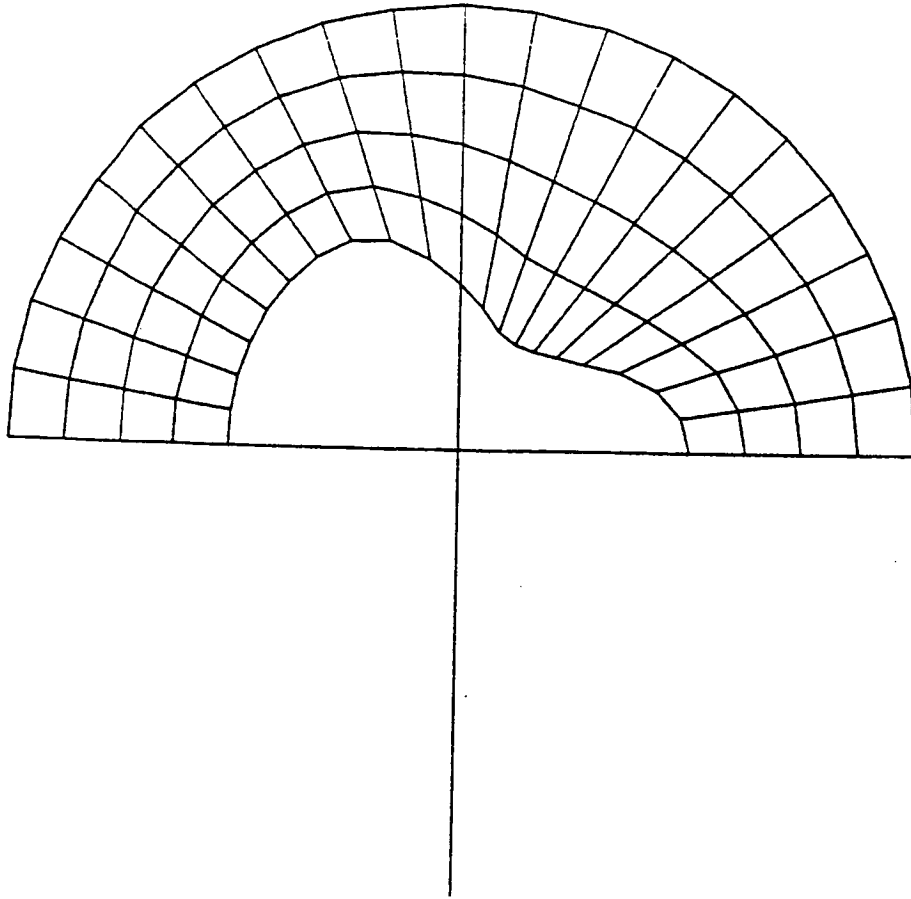
Figure 5--Karman-Trefftz Grid Construction
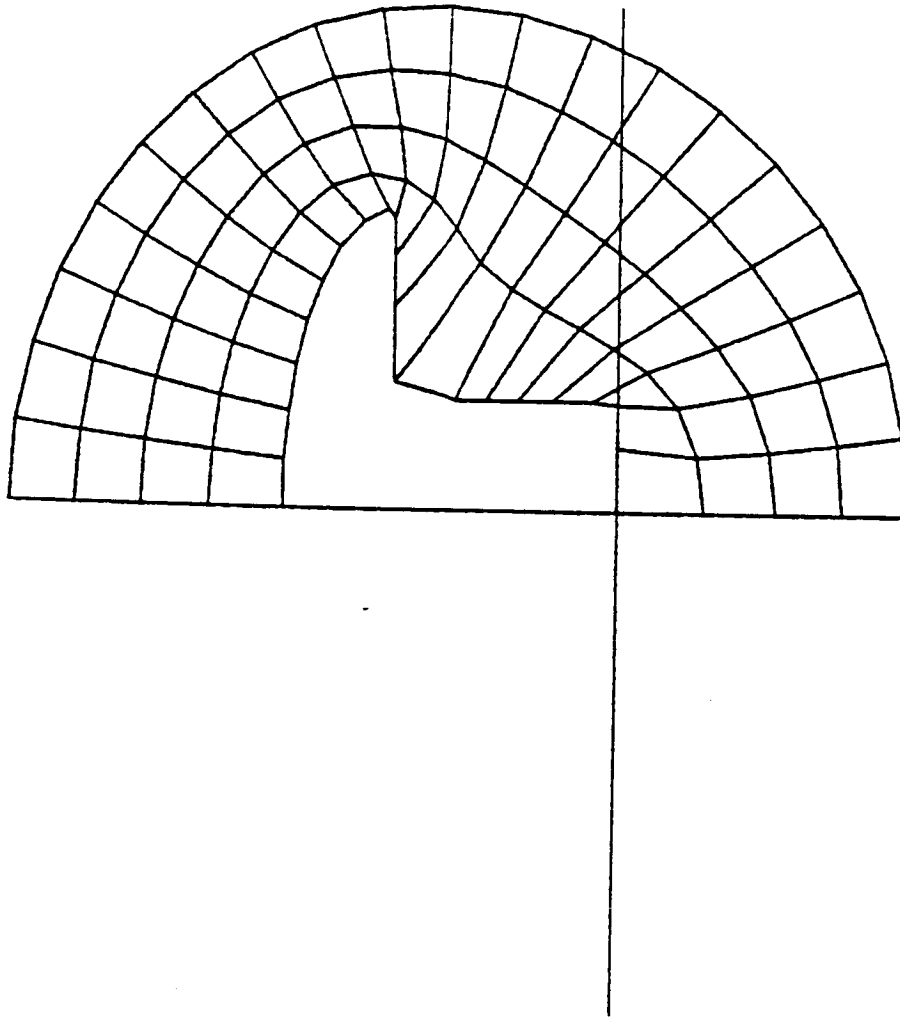
Figure 6--Grid in Physical Plane, Coarse Mesh
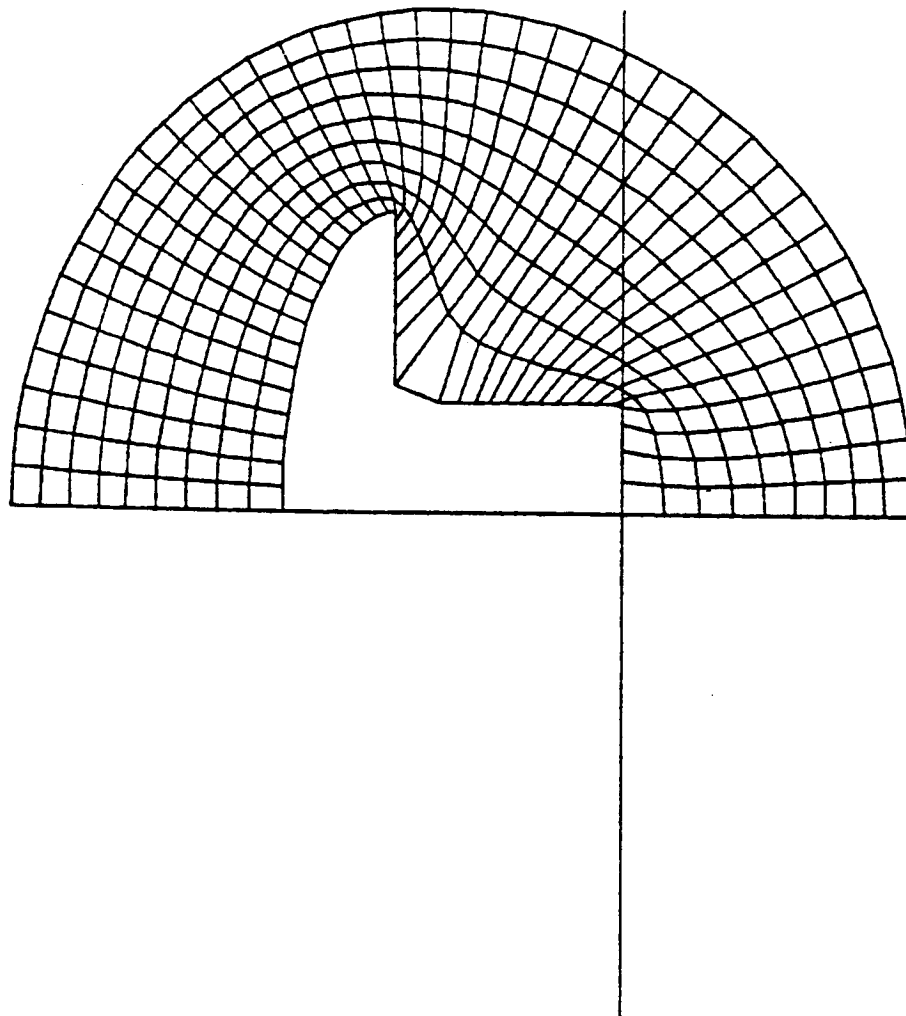
Figure 7--Grid in Physical Plane, Intermediate Mesh
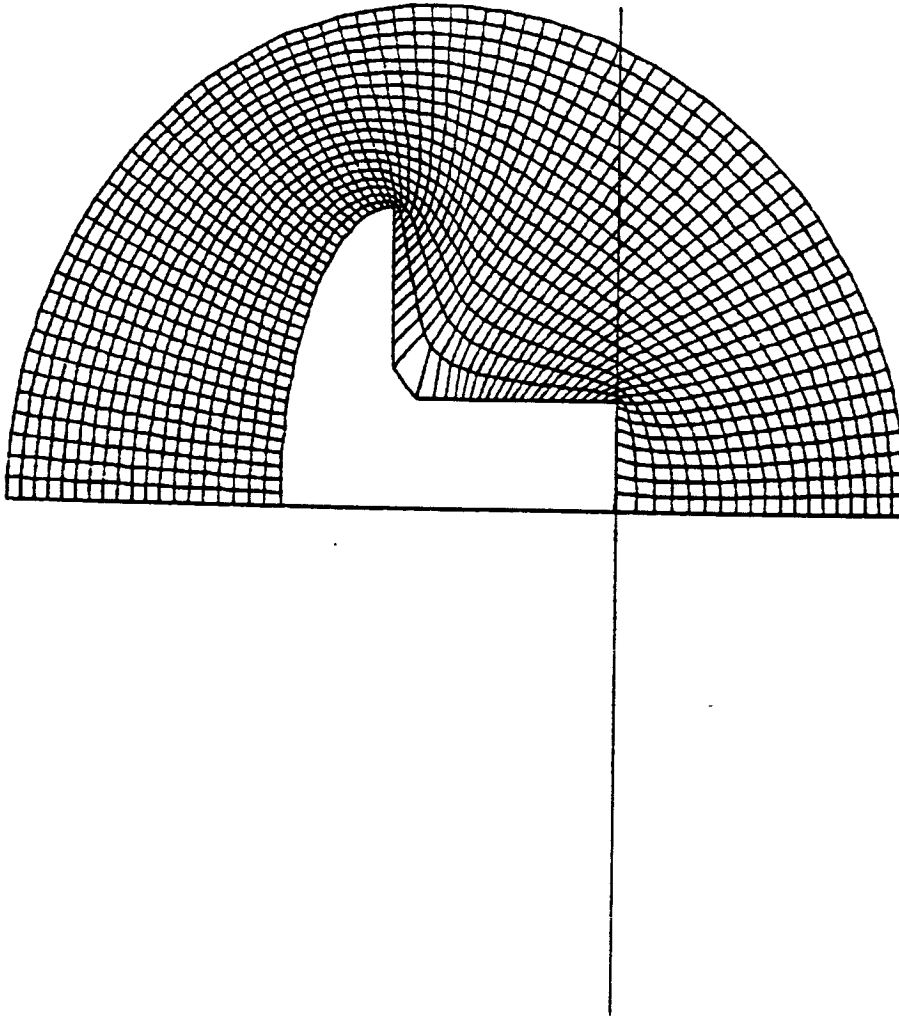
Figure 8--Grid in Physical Plane, Fine Mesh

3. Schwarz-Christoffel Transformation

The basic idea for an alternative transformation is presented by Hall (3), who refers to it as a "point-wise Schwarz-Christoffel transformation." Strictly speaking, however, the transformation is simply a power-law transformation taken from elementary complex number theory (Spiegel (4)). Motivation for the transformation is presented in Figure 9, where it is seen that points on the positive real axis in the physical plane remain on the real axis in the mapped plane, while points lying on a ray at angle in the physical plane are mapped onto the negative real axis in the mapped plane. Repeated applications of the transformation to every point on the physical body allows the mapping of the physical body onto the real axis. Finally, a grid can be generated in the mapped plane and mapped back into physical space using the inverse transformations. However, unlike the Karman-Trefftz transformation, the utilization of polar coordinates in the mapped plane for grid generation was found to provide inferior grid properties compared with the use of Cartesian coordinates in the mapped plane, because the axes in the mapped space do not correspond to $\theta=0$ and in the physical space. The following figures show that the grid generated is of no practical use. A listing of the computer program utilized for the Schwarz-Christoffel transformation is given in Appendix 2.

Figure 10 shows the Aerobrake body in its initial orientation. It has been rotated with respect to Figure 1 so that the exponent in Figure 9 will be finite. The computer program then calculates the angle which a line from point 0 to point 1 makes with the real axis, and calculates the exponent for the first transformation. This transformation results in the moving of point 1 to the real axis, as shown in Figure 11. Note that point 2 has been "lost" in Figure 11, due to interpolation. A more refined representation of the body shows that this error can be made arbitrarily small. Alternatively, an improved interpolation routine will eliminate this problem altogether. Next, point 2 is brought up to the real axis, as shown in Figure 12, followed by the positioning of point 3 on the real axis, as shown in Figure 13. Finally, point 4 is brought up to the real axis by yet another transformation, whereupon the Kutta-Joukowsky transformation is applied to map the transformed body onto a near-circle, as shown in Figure 14. Now equi-angular radial lines and their orthogonal complements are constructed, as shown in Figure 15. Now, however, when the resulting grid is mapped back into the physical plane, the grid is seen to overlap into the lower half plane, as shown in Figure 16. This is clearly an unacceptable grid for computation.

14-14
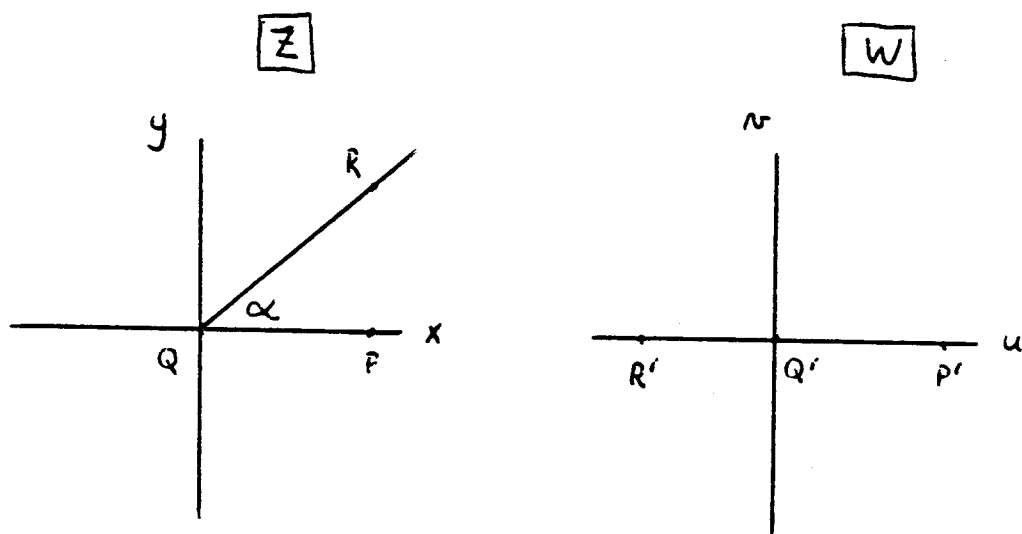
$$W = z^{\pi/\alpha}$$

Figure 9--Power Law Transformation
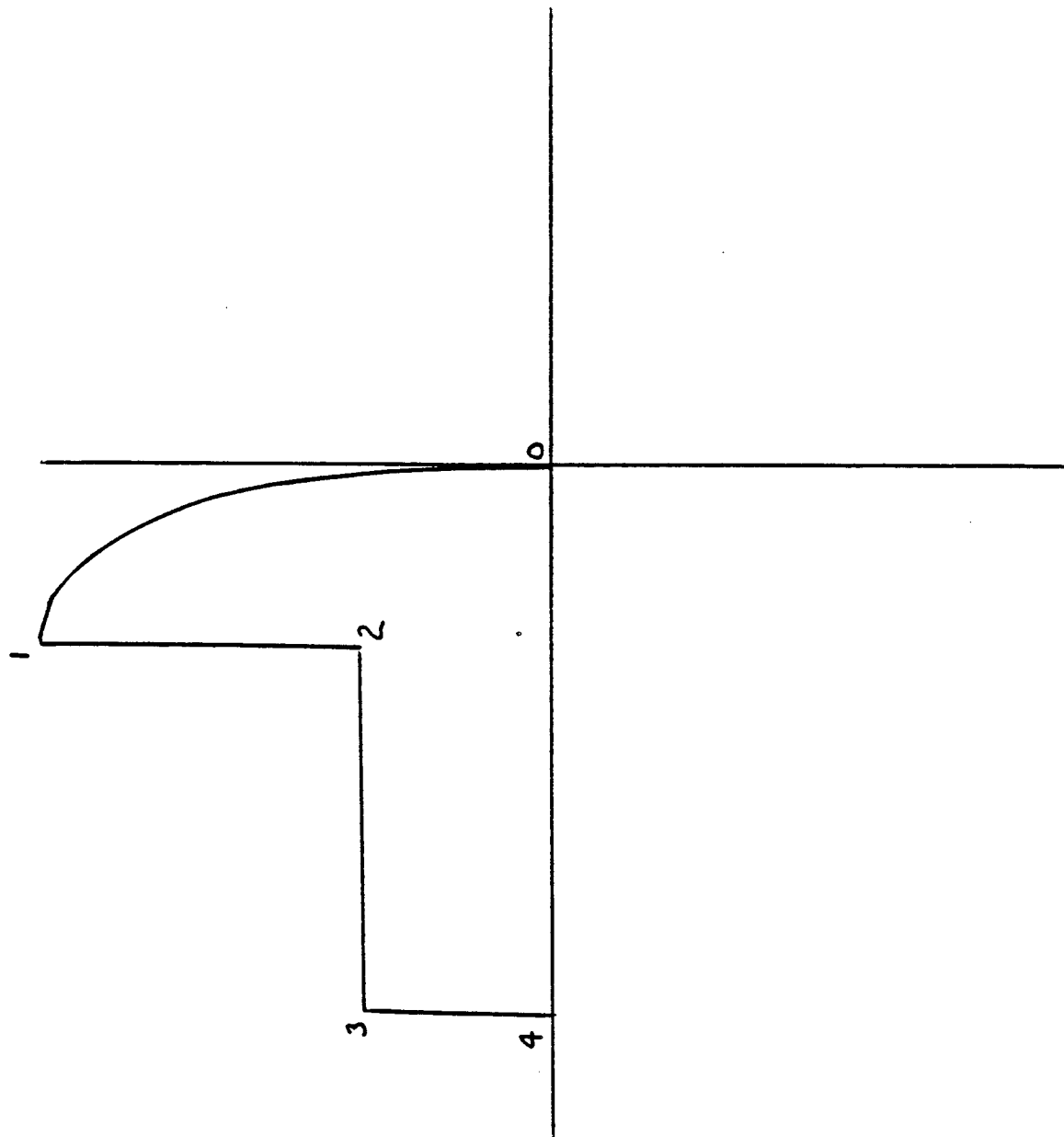
Figure 10--Mushroom Orientation for Point-Wise Schwarz-Christoffel Transformation
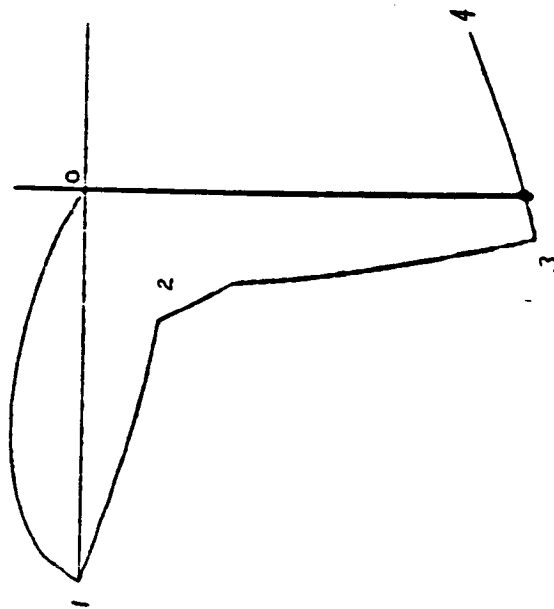
Figure 11--Result of First Transformation
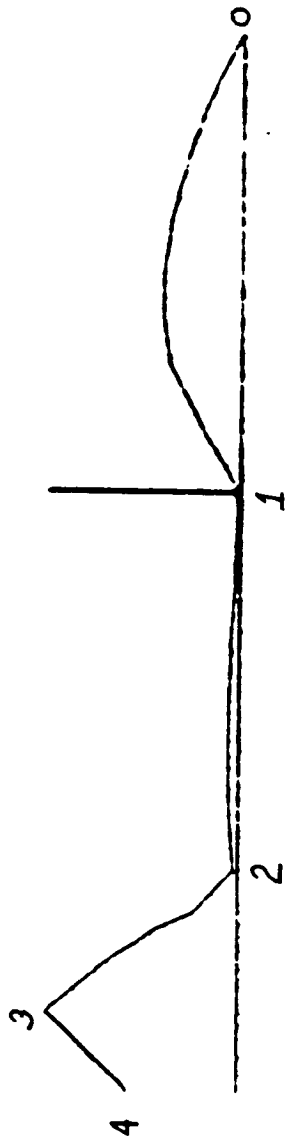
Figure 12--Result of Second Transformation

Figure 13--Result of Third Transformation
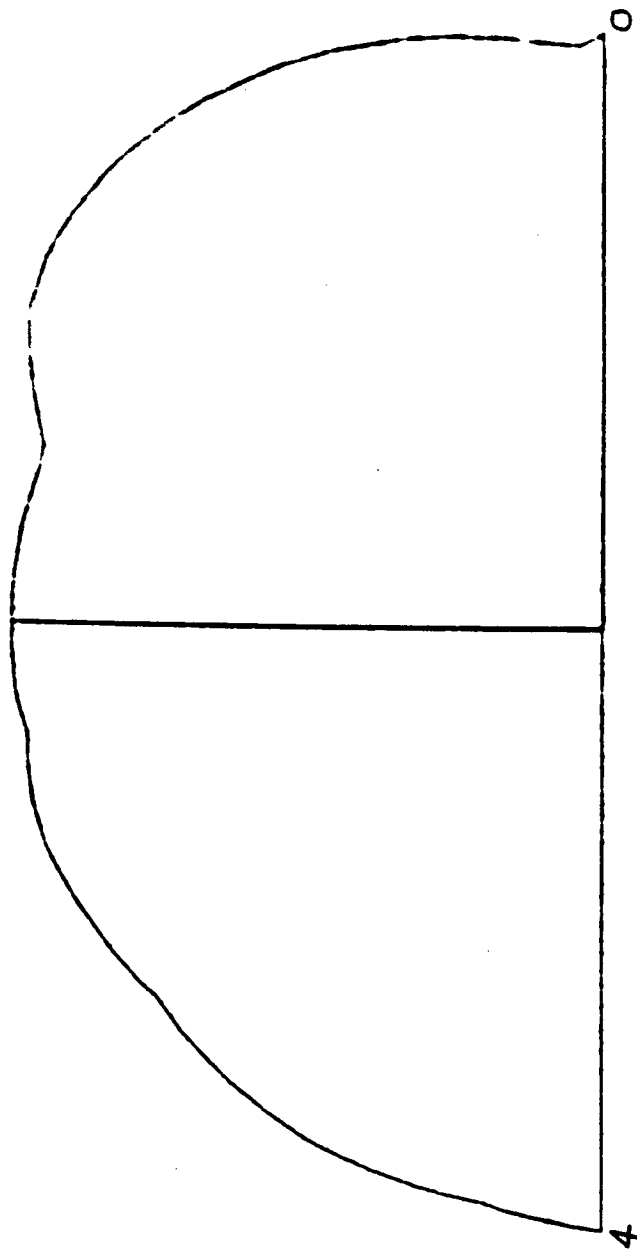
FIG. 14 -- After Applying Kutta-Joukowsky Transformation)
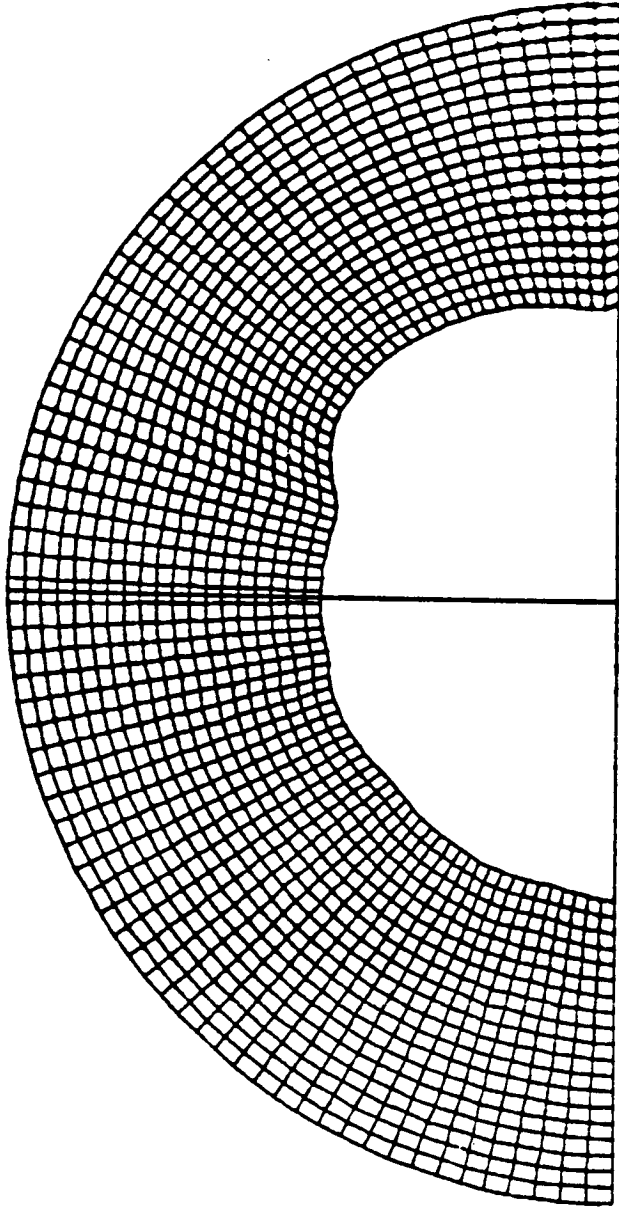
Figure 15--Grid Construction in Mapped Plane

DEN= 0.20915E+01
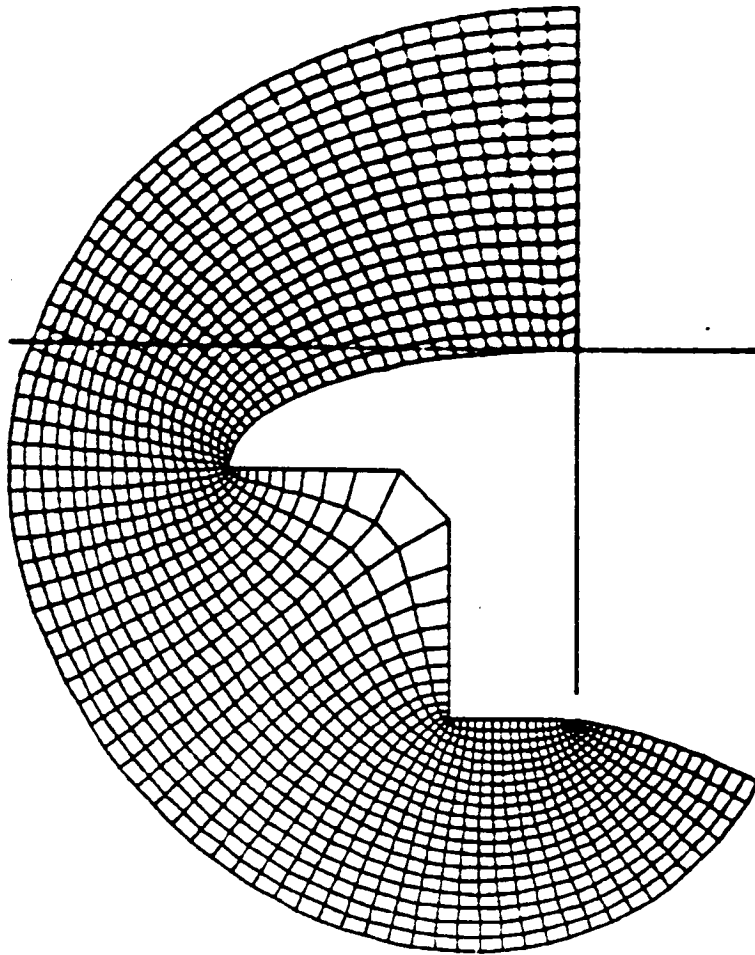
Figure 16--Grid in Physical Plane

DEN- ● 16236E+01

It is noted that the creation of a grid in the mapped plane utilizing polar coordinates fails to generate an acceptable grid in the physical plane due to the covering of more than the upper half plane in the latter by the upper half plane in the mapped plane. In order to overcome this shortcoming, various schemes were tried, including the addition of a tail to the mushroom in the physical plane, also the utilization of a quarter circle in the mapped plane instead of a semi-circle. The results of these attempts are presented in Figures 17-19. The introduction of a singularity downstream of the mushroom is evident. It became evident that the intermediate utilization of the Kutta-Joukowsky transformation to map the real axis onto a circle is not appropriate for the point-wise Schwarz-Christoffel transformation. Rather, the grid is to be generated by Cartesian coordinates in the mapped plane, as shown in Figures 20 and 21.
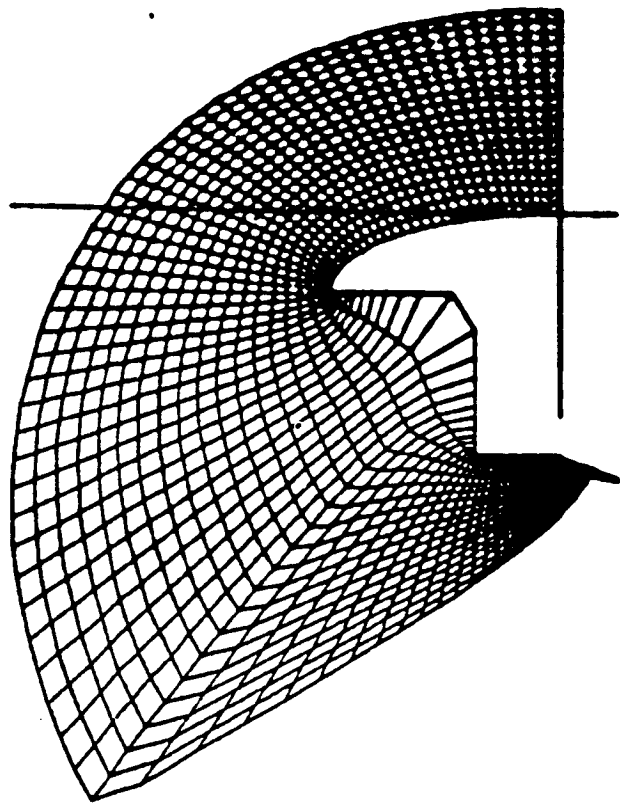
Figure 17 - Quarter Circle, No Tail
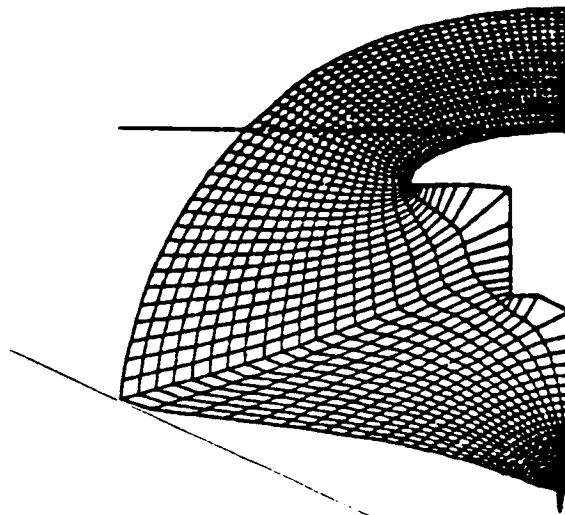
FIGURE 18- QUARTER CIRCLE, TAIL
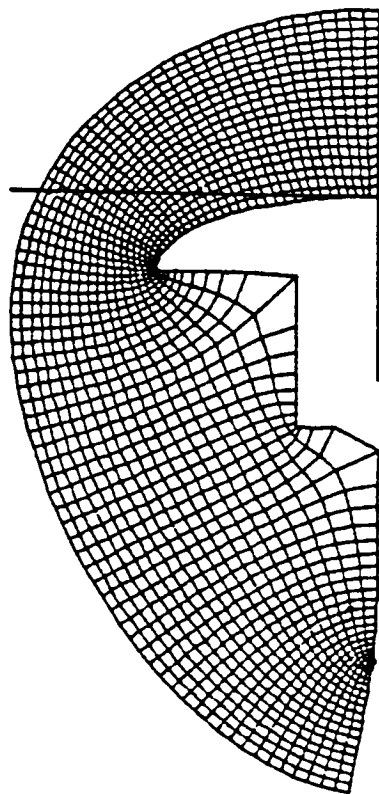
FIGURE 19 — SEMI CIRCLE, TAIL
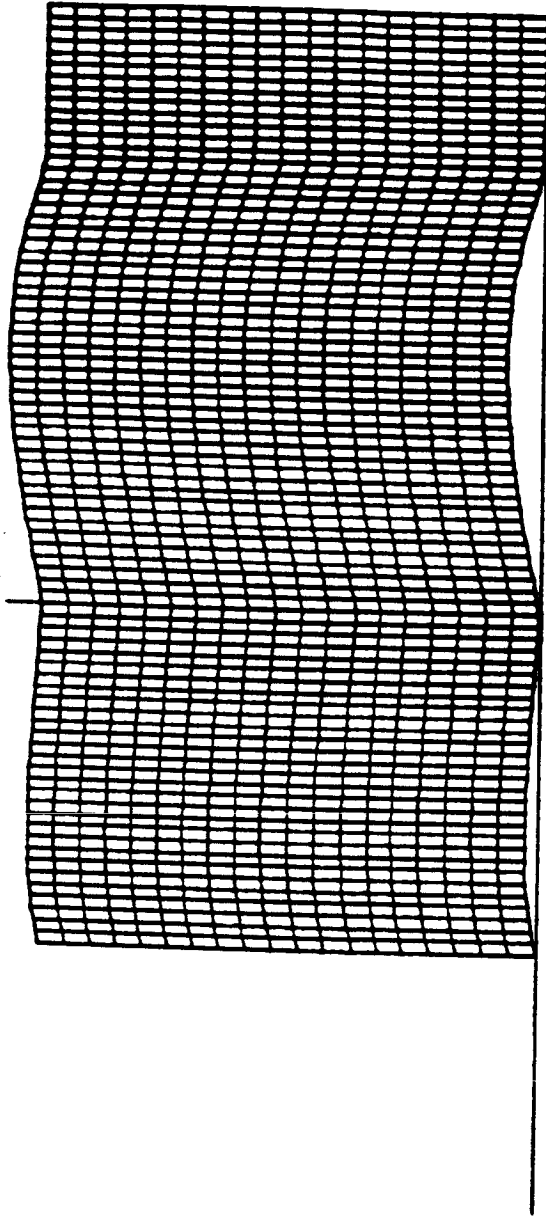
Figure 20--Cartesian Grid Construction in Mapped Plane

DEN= 0.18047E+01
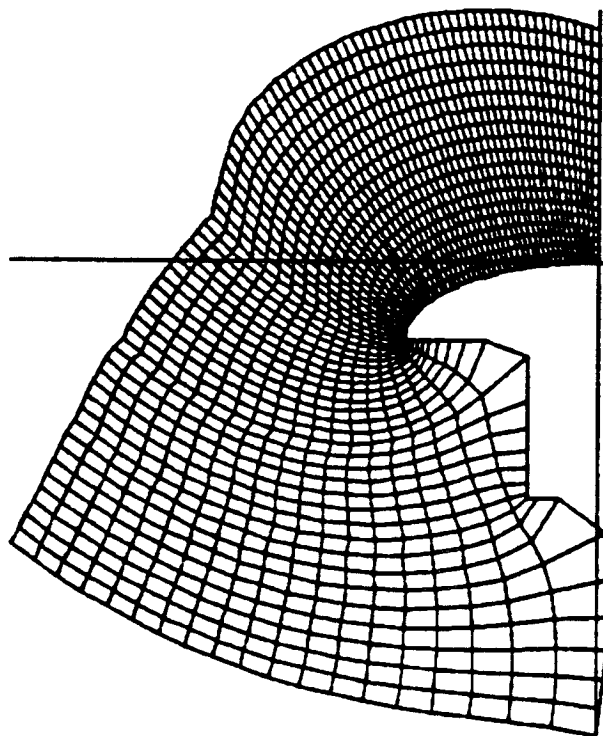
FIGURE 21-TAIL, CARTESIAN COORDINATES

4. Transformation of a Square

As noted earlier, it was desired to transform a simpler shape in order to gain insight into the transformations. A square was chosen for this effort, and the results are presented in the following figures. Trends similar to those observed for the mushroom are noted.

Figure 22 shows the square after the first Karman-Trefftz transformation has been applied. One corner has been smoothed out, and one remains. Figure 23 shows the square after the second corner has been removed. It is this latter shape which is utilized in Figure 24 for the construction of a polar grid, similar to what was done in Figure 5 for the Aerobrake body. Figure 25 shows the grid after the first inverse transformation, and Figure 26 shows the final grid in physical space.

The original square is shown in Figure 27, with the point-numbering convention for the Point-wise Schwarz-Christoffel transformation. After moving point 1 down to the real axis, the square takes on the shape shown in Figure 28. The next transformation moves point 2 up to the real axis, as shown in Figure 29; followed by the positioning of point 3 on the real axis as shown in Figure 30. Again the Kutta-Joukowsky transformation is applied, giving the near-circle shown in Figure 31. The polar grid is constructed in Figure 32, and the first inverse transformation yields the grid shown in Figure 33. However, Figure 34, which displays the result of the second inverse transformation, indicates trouble in that the grid begins to overlap itself. Figure 35, the grid in physical space, shows that, although the square has been successfully mapped back into physical space, the grid has not fared so well. Apparently, Figure 33 contains the explanation for the failure. Part of the grid in Figure 33 lies at an angle with the real axis which is greater (More negative) than that of line segment 2-3. It is this portion of the grid which cannot be successfully mapped back into physical space, probably because it is on another branch. Again, a Cartesian coordinate system is constructed in Figure 36, corresponding to Figure 20 for the Aerobrake body. When this is mapped back into physical spcae, Figure 37 results. The odd shape of this grid at infinity is clearly unacceptable. Instead, a tail is added to the square, both upstream and downstream, and the results are shown in Figures 38 and 39. The rear side of the square has been "lost" again, due to the interpolation scheme, as can be seen by comparing Figures 38 and 39.

KARMAN TREFFTZ
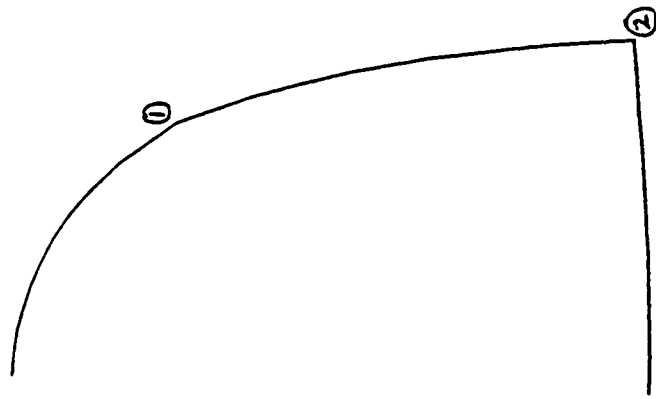
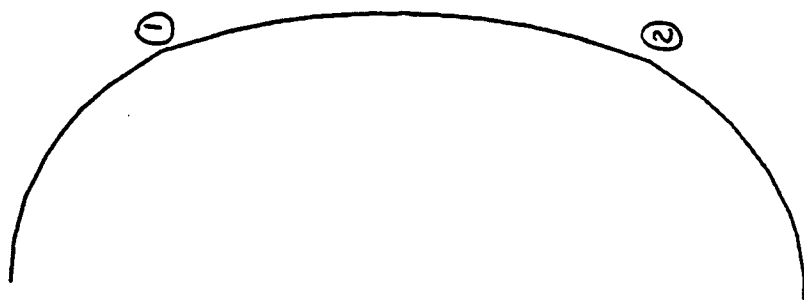Figure 22--Square after First Karman-Trefftz Transformation

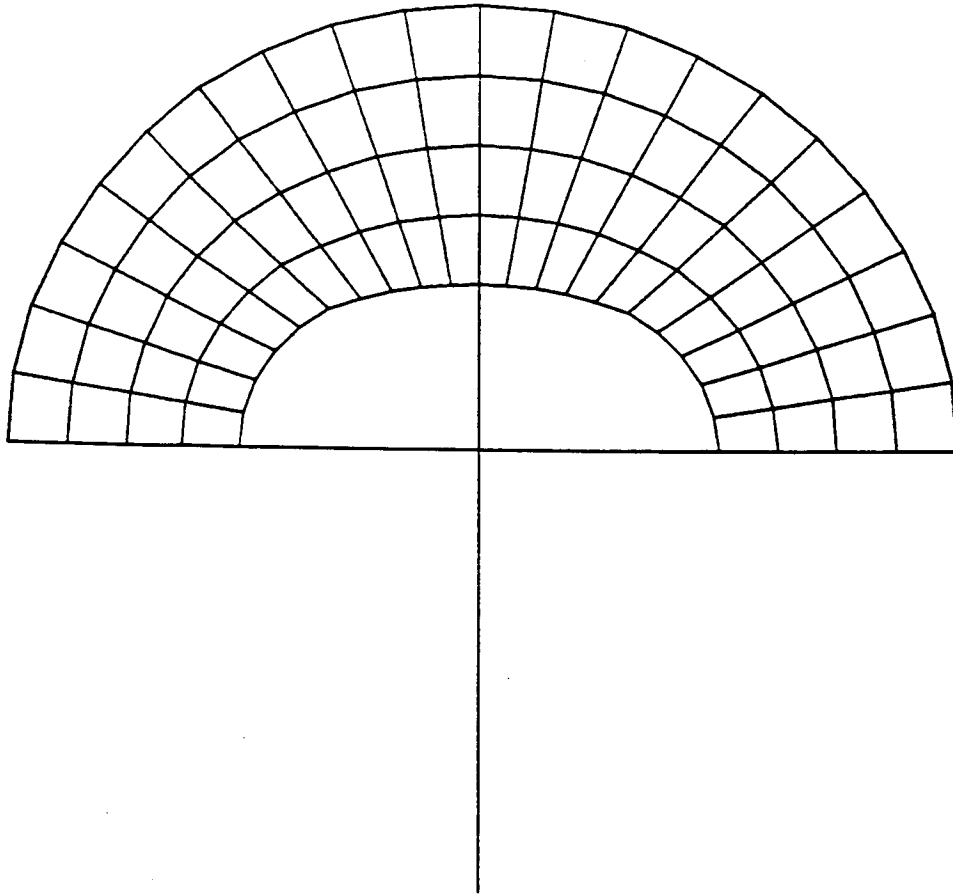Figure 23--Result of Second Transformation
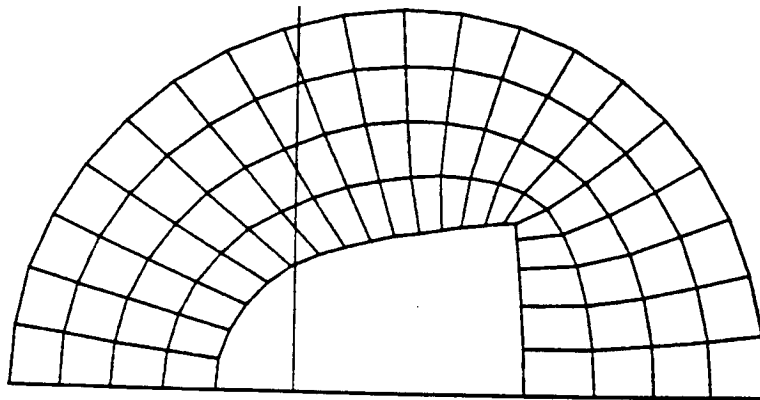
Figure 24--Grid Construction in Mapped Plane

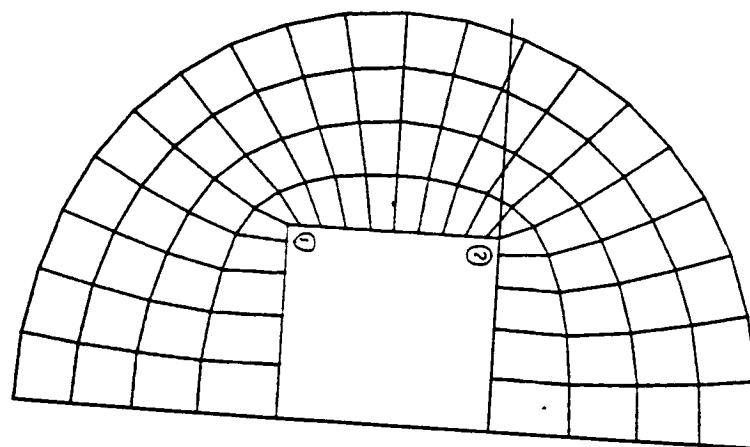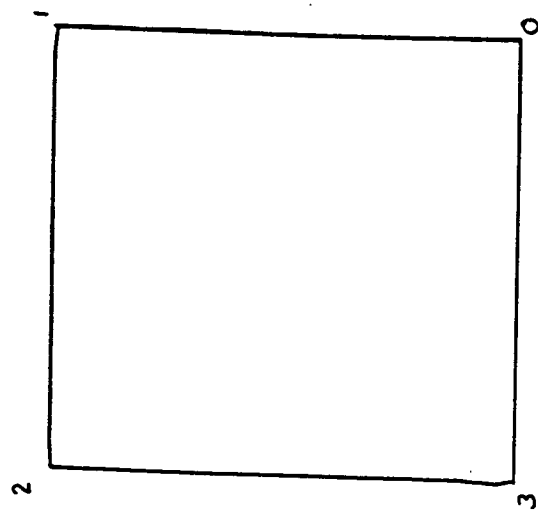Figure 25--Grid after First Inverse Transformation

Figure 26--Grid in Physical Space

POINT-WISE SCHWARZ-CHRISTOFFEL

Figure 27--Orientation of Square for Point-Wise Schwarz-Christoffel
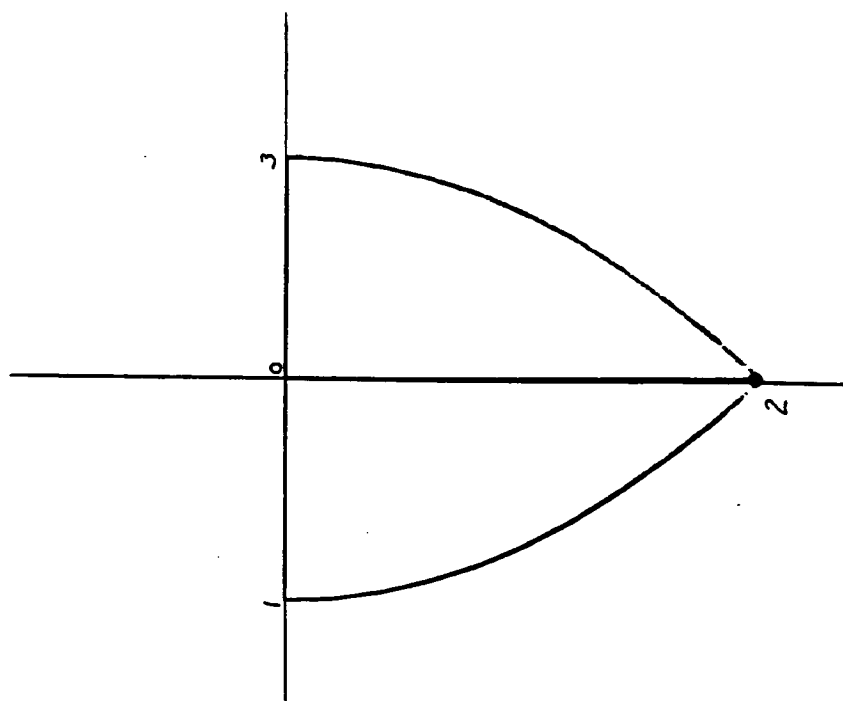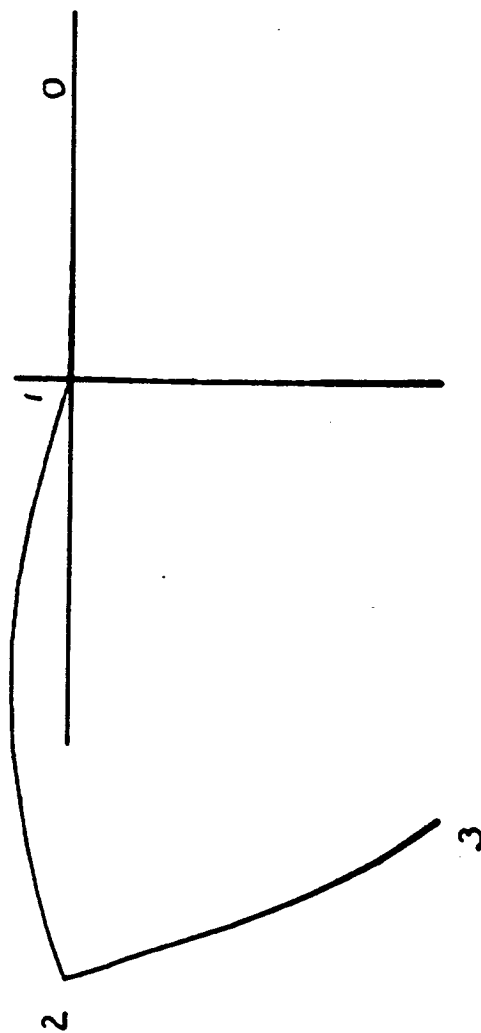Transformation

Figure 28--Result of First Transformation

14-36

Figure 29--Result of Second Transformation

Figure 30--Result of Third Transformation
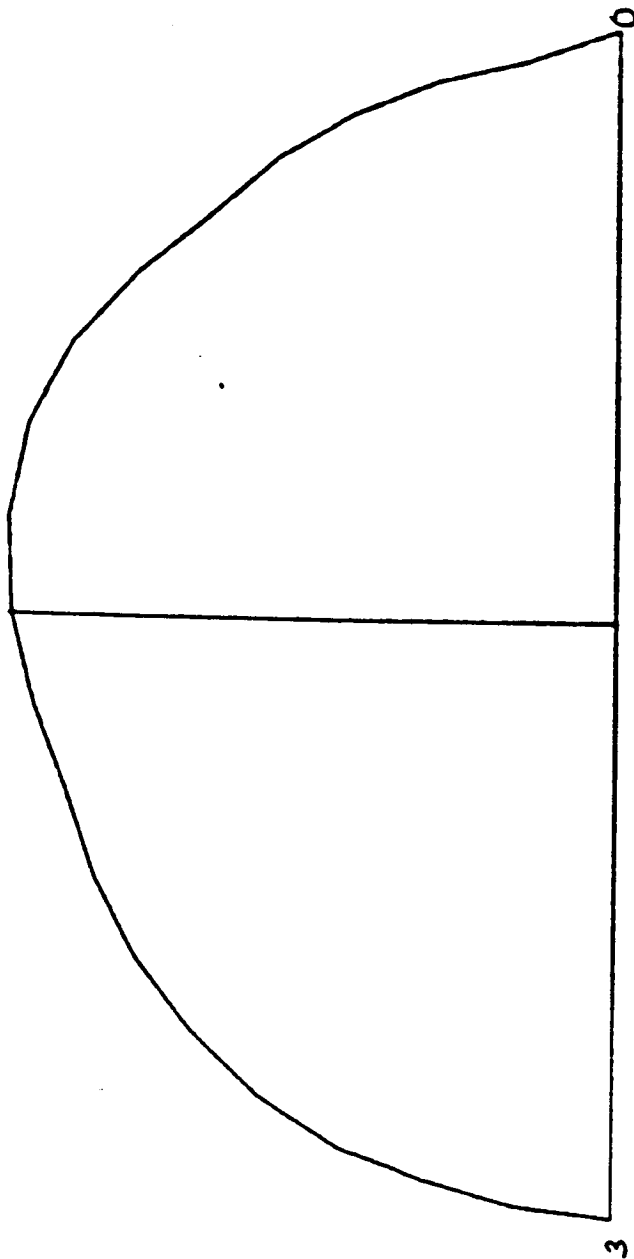
FIGURE 31 - After Kutta-Joukowsky Transformation
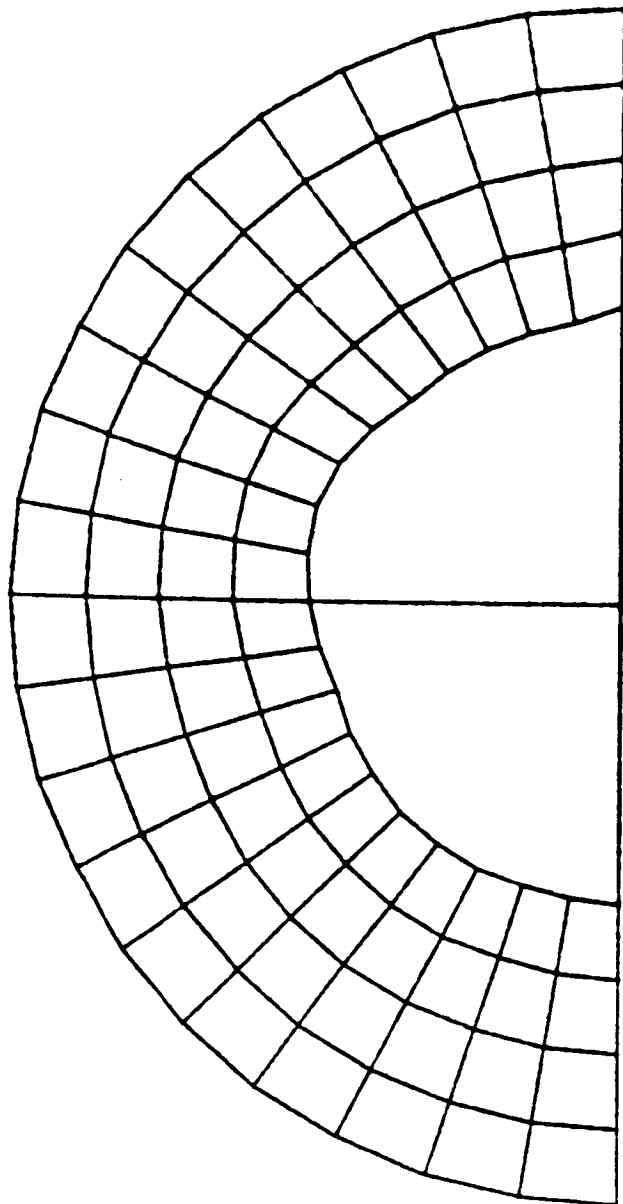
Figure 32--Polar Grid Generation
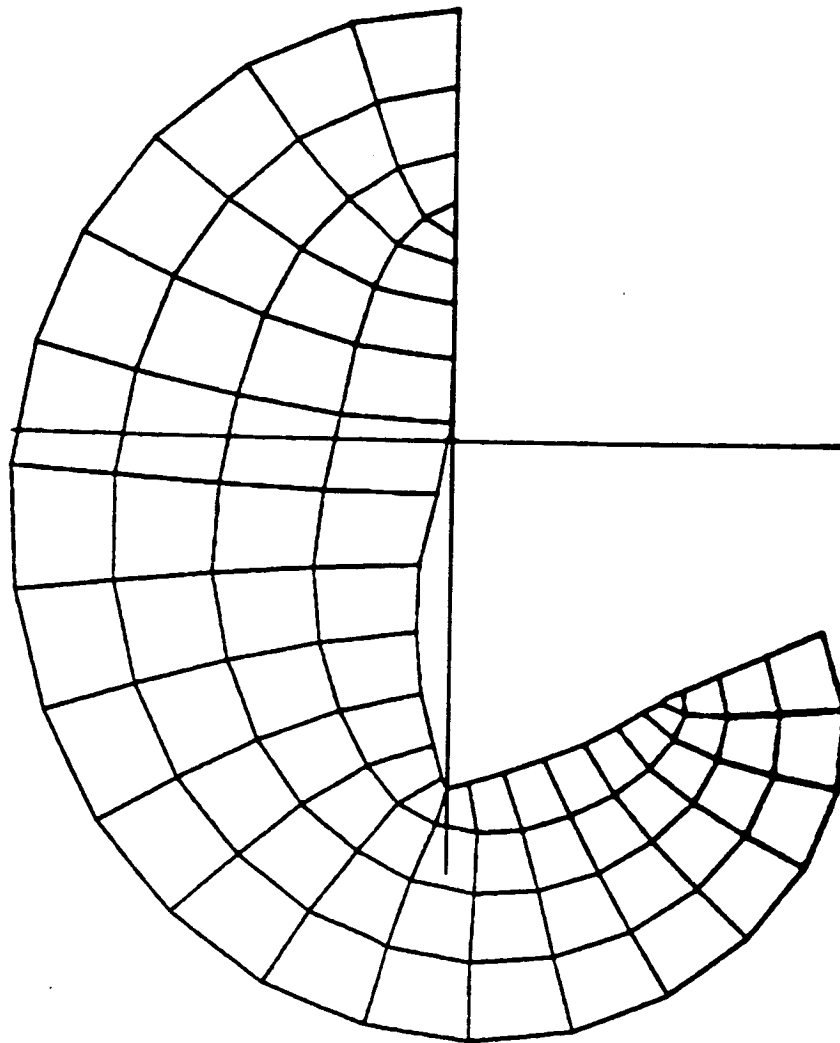
Figure 33--Polar Grid after First Inverse Transformation

14-41
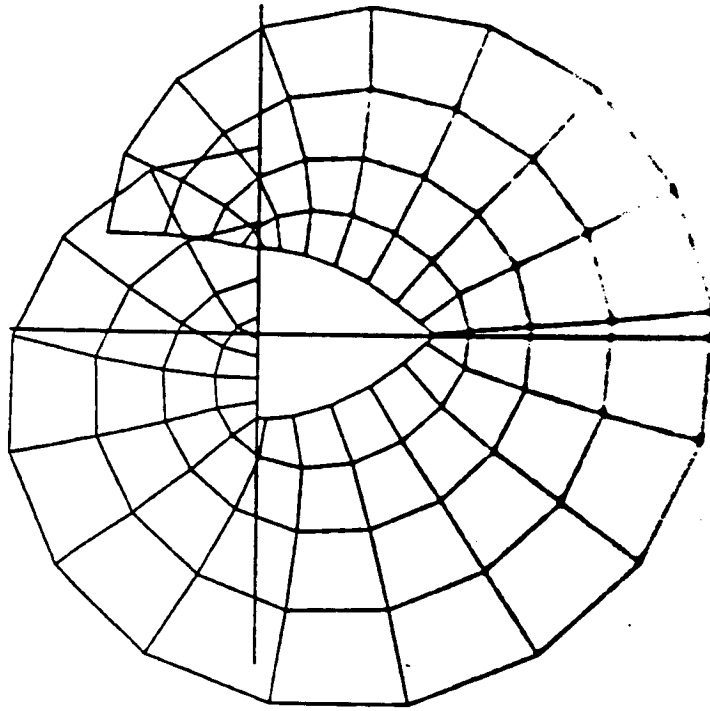
Figure 34--Polar Grid after Second Inverse Transformation
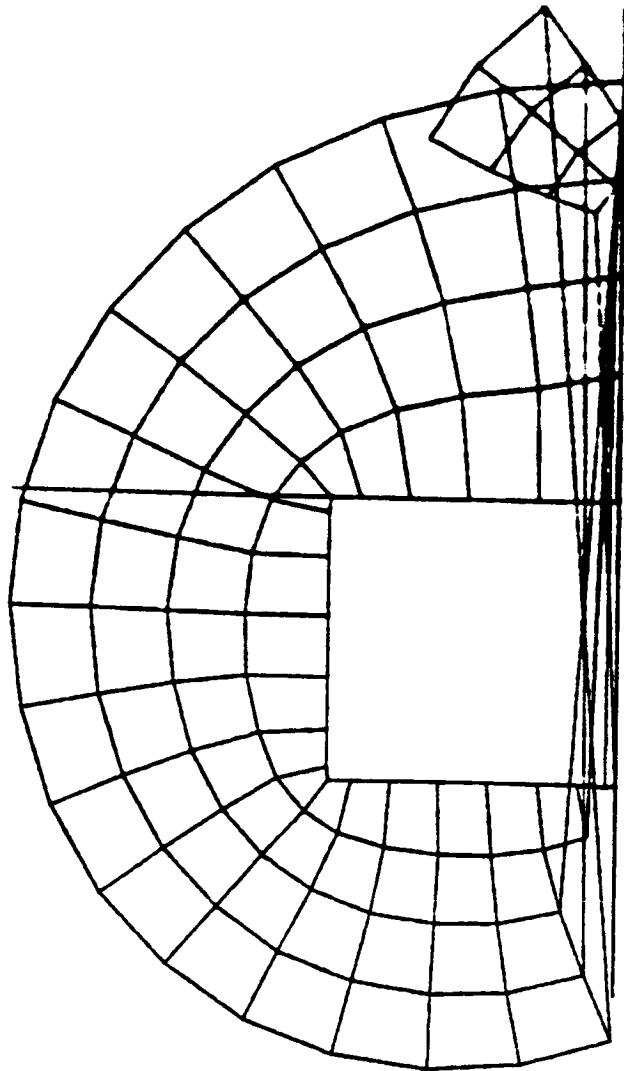
Figure 35--Polar Grid in Physical Space

DEN= 0.18047E+01
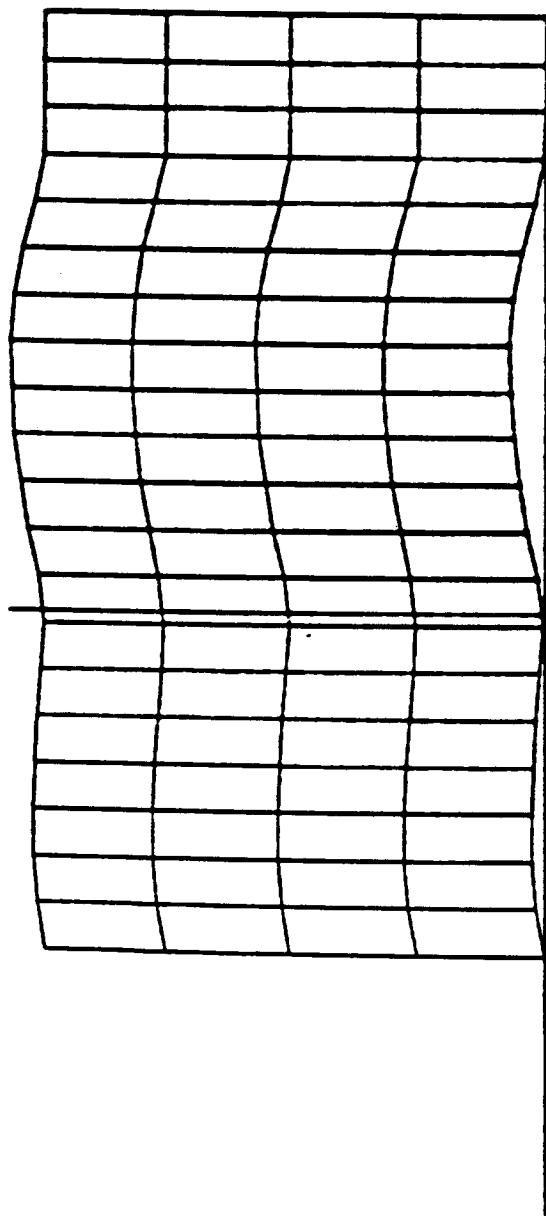
Figure 36--Cartesian Grid Generation in Mapped Space

DEN= 0.29321E+01
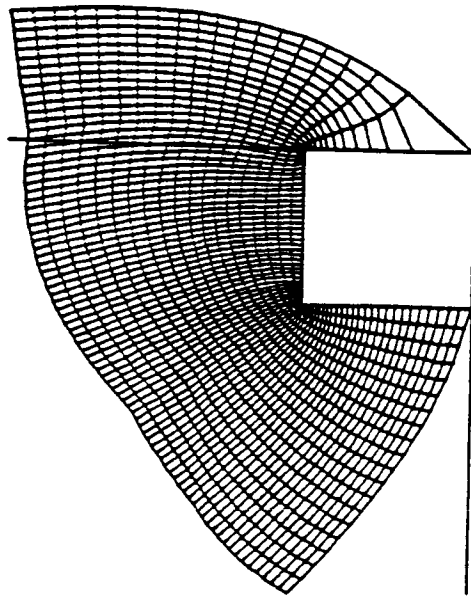
Figure 37--Cartesian Grid in Physical Plane
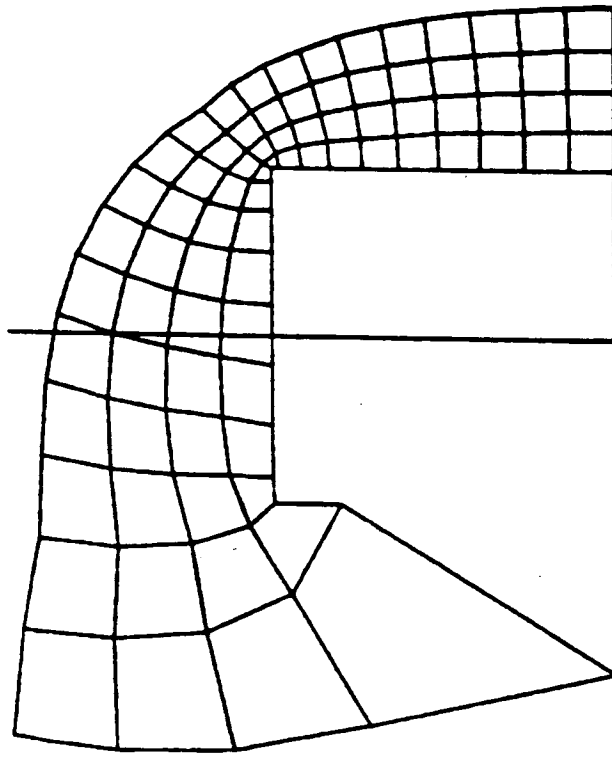
DEN= 0.17915E+01

Figure 38--Cartesian Grid for Square with Tails in Physical Space,
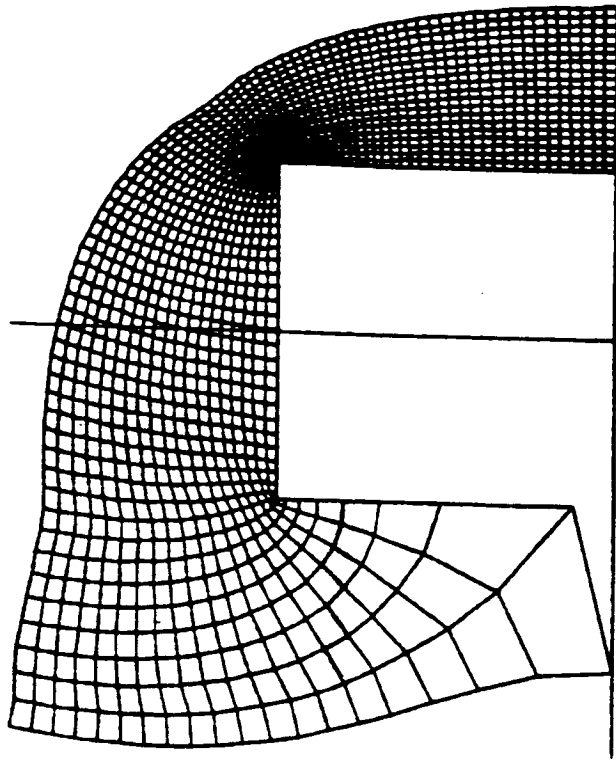Coarse Mesh

DEN= 0.17912E+01

Figure 39--Cartesian Grid for Square with Tails, Fine Mesh

5. Conclusions and Recommendations

Table 1 summarizes the advantages and disadvantages of the two transformations. It is seen that both have complementary advantages and disadvantages. One disadvantage of both transformations which is listed in Table 1 refers to poor resolution at concave corners. This is not a serious disadvantage, at least for the flow past the Aerobrake Body, because the flow in this region is of lesser interest.

Based on the results, the following recommendations are offered:

1. The Karman-Trefftz transformation appears to be best suited to finite bodies around which the entire flow field is desired. The type of mesh which results under this transformation is an "O" mesh.

2. The point-wise Schwarz-Christoffel transformation appears to be best suited to infinite bodies, or finite bodies with long trailing wakes. The type of mesh results under this transformation is a "C" mesh.

Transformation	Advantages	Disadvantages
Karman-Trefftz	1. Clustering at Convex Corners 2. Orthogonal at Front & Rear 3. Well-Shaped at Infinity	1. Not Orthogonal on Stem 2. Poor Resolution at Concave Corners
Point-Wise Schwarz-Christoffel	1. Clustering at Convex Corners 2. Orthogonal Everywhere	1. Odd-Shaped at Infinity 2. Poor Resolution at Concave Corners

Table 1--Comparison of the Two Transformations

6. References

1. Li, C.P., 1985, Numerical Procedure for Three-Dimensional Hypersonic Viscous Flow over Aerobrake Configuration, NASA TM58269.

2. Moretti, G., 1976, Conformal Mappings for Computations of Steady, Three-Dimensional, Supersonic Flows, Numerical/Laboratory Comp. Methods in Fl. Mech., A.A. Pouring and V.I. Shah, Ed., ASME, pp. 13-28.

3. Hall, D.W., 1980, A Three-Dimensional Body-Fitted Coordinate System for Flow Field Calculations on Asymmetric Nosetips, Numerical Grid Generation Techniques, NASA Conf. Publ. 2166.

4. Spiegel, M.R., 1964, Complex Variables, McGraw-Hill.

1. Report No. NASA CR 171984	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle NASA/ASEE Summer Faculty Fellowship Program--1986 Volume 1		5. Report Date June 1987
		6. Performing Organization Code
7. Author(s) Editors: Bayliss McInnis and Stanley Goldstein		8. Performing Organization Report No.
		10. Work Unit No.
9. Performing Organization Name and Address The University of Houston--University Park and Texas A&M University		11. Contract or Grant No. NGT-44-005-803
		13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The Johnson Space Center (JSC) NASA/ASEE Summer Faculty Fellowship Program was conducted by the University of Houston--University Park and the Johnson Space Center. The ten week program was operated under the auspices of the American Society for Engineering Education (ASEE). The program at JSC, as well as the programs at other NASA Centers, was funded by the Office of University Affairs, NASA Headquarters, Washington, D.C. The basic objectives of the programs, which began in 1965 at JSC and in 1964 nationally, are (a) to further the professional knowledge of qualified engineering and science faculty members; (b) to stimulate an exchange of ideas between participants and NASA; (c) to enrich and refresh the research and teaching activities of participants' institutions; and (d) to contribute to the research objectives of the NASA Centers.

Each faculty fellow spent ten weeks at JSC engaged in a research project commensurate with his interests and background and worked in collaboration with a NASA/JSC colleague. This document is a compilation of the final reports on the research projects done by the faculty fellows during the summer of 1986. Volume 1 contains sections 1 through 14, and volume 2 contains sections 15 through 30.

17. Key Words (Suggested by Author(s))	18. Distribution Statement Unclassified - Unlimited

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 359	22. Price* NTIS